



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 157842

TO: Jon E Angell
Location: rem/2D20/2C18
Art Unit: 1635
Friday, July 01, 2005

Case Serial Number: 09/945173

From: Noble Jarrell
Location: Biotech-Chem Library
Rem 1B71
Phone: 272-2556

Noble.jarrell@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

157842

me

From: Chan, Christina
Sent: Tuesday, June 28, 2005 5:24 PM
To: Angell, Jon E; STIC-Biotech/ChemLib
Subject: RE: RUSH Sequence Search Request 09/945,173

Please ~~rush~~. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

RECEIVED
JUN 29 2005
STIC

-----Original Message-----

From: Angell, Jon E
Sent: Tuesday, June 28, 2005 5:19 PM
To: Chan, Christina
Subject: RUSH Sequence Search Request 09/945,173

Chrisitina,

I would like to request a RUSH search as indicated below. The cases may be allowable and I need the interference search ASAP.

If approved please forward the request to STIC and cc me.

Thanks,

Eric

SEARCH REQUEST FORM
Scientific and Technical Information Center

Examiner# : 78697
Art Unit : 1635
Phone Number: 571-272-0756
Date: 6/28/05
Serial Number: 09/945,173 (Meyers)
Mailbox & Bldg/Room Location: REMSEN 2C18
Results Format Preferred (circle): Paper

I would like to have a standard and interference search performed using the following SEQ. ID NOs. from application : 09/945,173

SEQ ID NO. 1 (nucleic acid 1694 nucleotides in length)

STAFF USE ONLY

Searcher: noble
Searcher Phone: 2-
Date Searcher Picked up: 7/1/05
Date Completed: 7/1/05
Searcher Prep/Rev. Time: 5
Online Time: 3

Type of Search

NA#: 2 AA#: 1
Interference: ✓ SPDI:
S/L: Oligomer:
Encode/Transl:
Structure#: Text:
Inventor: Litigation:

Vendors and cost where applicable

STN:
DIALOG:
QUESTEL/ORBIT:
LEXIS/NEXIS:
SEQUENCE SYSTEM: Compu
WWW/Internet:
Other(Specify):

THIS PAGE BLANK (USPTO)

SEQ ID NO. 2 (protein 236 amino acids in length)
SSE ID NO: 3 (nucleic acid 711 nucleotides in length)

NOTE: SEQ ID NO: 1 is a full length DNA
SEQ ID NO: 3 is the only the coding region of SEQ ID NO. 1
SEQ ID NO: 2 is a protein encoded by SEQ ID NO: 1 and 3

Please feel free to contact me if you need any help...

Thanks for your help,

Jon Eric Angell
Art Unit 1635
Office: REMSEN 2D20
mailbox: REM 2C18
571-272-0756

(STIC)

JUN 29 2000

RECEIVED

MEJ

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 13:28:23 ; Search time 6313.94 Seconds
(without alignments)
10212.469 Million cell updates/sec

Title: US-09-945-173-1

Perfect score: 1694

Sequence: 1 agggaggcagtggaagat.....aaacatgaattattgaacta 1694

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gss1:*
- 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1621.2	95.7	1867	3	CR619977	full-length
2	880	51.9	2013	3	AK028195	Mus muscu
3	879	51.9	918	5	BX432920	BX432920
4	868.6	51.3	2006	3	AK028269	Mus muscu
5	799	47.2	912	5	BX348224	BX348224
6	791.8	46.7	797	1	AL516309	AL516309
7	791.8	46.7	918	4	BI862084	603390970
8	770.4	45.5	1961	3	AK011280	Mus muscu
9	750.6	44.3	815	5	BQ215951	AGENCOURT
10	750	44.3	875	7	CK456385	917058 MA
11	748.2	44.2	950	5	BQ893351	AGENCOURT
12	747	44.1	855	7	CK455329	915906 MA
13	718.4	42.4	720	7	CN361073	328775509
14	718.4	42.4	718	5	BX099226	BX099226
15	712.4	42.1	989	4	BM451532	AGENCOURT
16	707.4	41.8	732	4	BI333788	602999279
17	697	41.1	881	5	BU178665	AGENCOURT
18	691.2	40.8	720	4	BG501635	602548678
19	686	40.5	1118	1	AL517896	AL517896
20	685.6	40.5	699	7	CN361072	170006000
21	674	39.8	785	4	BI759699	603045561
22	669.2	39.5	933	3	AK016099	Mus muscu
23	664.2	39.2	902	4	BG258348	602379905
24	662.8	39.1	959	4	BG031602	602299747

25	636.4	37.6	649	4	BG025782	602274775
26	625.6	36.9	696	7	CN794203	4129345 B
27	620.6	36.6	643	6	CA440480	UI-H-ED0-
28	619	36.5	986	4	BI334986	602998879
29	597.8	35.3	602	4	BG6113092	602640623
30	589.8	34.8	920	1	AA203636	2X58B31.1
31	588.8	34.8	827	2	BF571796	602078477
32	583	34.4	584	7	CN361075	170004241
33	574.2	33.9	706	7	CF730316	UI-M-GZ0-
34	567	33.5	579	4	BM674487	UI-E-E370-
35	564.4	33.3	832	1	AL516308	AL516308
36	562.6	33.2	665	4	BG080166	H3050803-
37	560.4	33.1	587	4	BM751558	K-EST0027
38	560	33.1	786	2	BF106997	601824352
39	559.6	33.0	831	4	BG400077	602442152
40	556.8	32.9	722	2	BB610950	BB610950
41	550.6	32.5	743	2	BB611463	BB611463
42	549.8	32.5	810	6	CD512097	AGENCOURT
43	544.4	32.1	577	5	BQ013770	UI-1-BC1p
44	540	31.9	551	1	AA662211	ns66d03.1
45	533.6	31.5	559	2	AW974777	EST386867

ALIGNMENTS

RESULT 1
LOCUS CR619977 1867 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DA005YF10 of Neuroblastoma of Homo sapiens (human).
ACCESSION CR619977
VERSION CR619977.1 GI:50500784
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1867)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1867)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
JOURNAL - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
source
1..1867
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA005YF10"
/issue_type="Neuroblastoma"
/plasmid="pCMVSPORT_6"

Query Match 95.7%; Score 1621.2; DB 3; Length 1867;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1623; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 62 CAGGTGTGGGAATCTTCGTTAGTCATCTCTATGCCAAATCAAGTCTGGGAATC 121
|||||
Db 1 CAGGTGTGGGAATCTTCGTTAGTCATCTCTATGCCAAATCAAGTCTGGGAATC 60

Qy	122	CATCATGGAACTGTGGGCTGCTCAGTGGATGT	CAGAGTT	CATGATT	CAAAAGAAAGGAACCC	181
Db	61	CATCATGGACTGTGGGCTGCTCAGTGGATGT	CAGAGTT	CATGATT	CAAAAGAAAGGAACCC	120
Qy	182	CAGAGAGAGAACCTACTACATAGAAATTATGGAGT	TTGGAGGCT	CTGTGGG	CAGATGCCA	241
Db	121	CAGAAGAGAGAACCTACTACATAGAAATTATGGAGT	TTGGAGGCT	CTGTGGG	CAGATGCCA	180
Qy	242	GCAGCGTGAAGAACACAAAGACAGTATTTCTACAACT	CCGTAAATGGT	TATTTTCGTAC	301	
Db	181	GCAGCGTGAAGAACACAAAGACAGTATTTCTACAACT	CCGTAAATGGT	TATTTTCGTAC	240	
Qy	302	ACGACTTAAACAAATAAGAAGTTCCTCCCAAACT	TTGCGTCCGT	CAATTTGGAAGCTCTCA	361	
Db	241	ACGACTTAAACAAATAAGAAGTTCCTCCCAAACT	TTGCGTCCGT	CAATTTGGAAGCTCTCA	300	
Qy	362	ACAGGGATTTGGTGGCCAACTCGGAGTCTTTGGTGA	CAAAATGGGGAT	TATATGATCAAGAACAGT	421	
Db	301	ACAGGGATTTGGTGGCCAACTCGGAGTCTTTGGTGA	CAAAATGGGGAT	TATATGATCAAGAACAGT	360	
Qy	422	TTGCTGATATAACCAATAACCACTGTTGGTAAATAGGGA	CTAAACTGGAC	CCAGATTCATGAAA	481	
Db	361	TTGCTGATATAACCAATAACCACTGTTGGTAAATAGGGA	CTAAACTGGAC	CCAGATTCATGAAA	420	
Qy	482	CAAAAGCCCATGAAGTTTAACTAGGACTGCTTTCTGGCT	GTGAGGATTTCAAT	CTCCAGAAG	541	
Db	421	CAAAAGCCCATGAAGTTTAACTAGGACTGCTTTCTGGCT	GTGAGGATTTCAAT	CTCCAGAAG	480	
Qy	542	AAATTAATTTGGACTGCACAAATCCACGGTACTTTAGCTG	CAGGTTCTTCAAT	TGCTGTGCA	601	
Db	481	AAATTAATTTGGACTGCACAAATCCACGGTACTTTAGCTG	CAGGTTCTTCAAT	TGCTGTGCA	540	
Qy	602	AGCTCAGTGTGTTTTTTGATAAGGTCATAGAGAAGATAC	TCTTTTAAAGAGAAGGTAATC	661		
Db	541	AGCTCAGTGTGTTTTTTGATAAGGTCATAGAGAAGATAC	TCTTTTAAAGAGAAGGTAATC	600		
Qy	662	AGATTCCAGGCTTTCCTGTATCGGAAAAAGATTTGGGG	CAGGAACATTTAAAGAGGCTTCATT	721		
Db	601	AGATTCCAGGCTTTCCTGTATCGGAAAAAGATTTGGGG	CAGGAACATTTAAAGAGGCTTCATT	660		
Qy	722	ATGACTGAATTTACACTCATCTTTGGAGAGGTGAGCAAG	CAGTGGCAGTTTTTTCACAGCT	781		
Db	661	ATGACTGAATTTACACTCATCTTTGGAGAGGTGAGCAAG	CAGTGGCAGTTTTTTCACAGCT	720		
Qy	782	CATCTTGCTGTGTTTCAATATTACCATCACAGCCCTTTT	TAAACAAATCATCTTAAAAATGCT	841		
Db	721	CATCTTGCTGTGTTTCAATATTACCATCACAGCCCTTT	TAAACAAATCATCTTAAAAATGCT	780		
Qy	842	ACCCTTCAGCCCTTACCCCTTTAAATGGAAAAATGAAAG	GAAGTGACAATACCGGAGGTCCAA	901		
Db	781	ACCCTTCAGCCCTTACCCCTTTAAATGGAAAAATGAAAG	GAAGTGACAATACCGGAGGTCCAA	840		
Qy	902	ACTTTGTCCCTGTGTTCTCTGTGTTCCCTTTACCTTT	CTGTCCTCTGTATAGATTTATGTAATA	961		
Db	841	ACTTTGTCCCTGTGTTCTCTGTGTTCCCTTTACCTTT	CTGTCCTCTGTATAGATTTATGTAATA	900		
Qy	962	GCCTTGTTGTAATAATGAGATGTTGTTCAAAATGATGCA	GTAAATGAGCAATGACAGTGTAC	1021		
Db	901	GCCTTGTTGTAATAATGAGATGTTGTTCAAAATGATGCA	GTAAATGAGCAATGACAGTGTAC	960		
Qy	1022	TGCAGAGAAAAATTTACTCTTGGCTTAGAACTGGAGGG	TTTTTATGGGTTCTGTAAATTTTCCC	1081		
Db	961	TGCAGAGAAAAATTTACTCTTGGCTTAGAACTGGAGGG	TTTTTATGGGTTCTGTAAATTTTCCC	1020		
Qy	1082	ACACTCATTTGCTGAAAGCTTAAATTAAGTACTTTCAAAA	ACGTTATCTCCATTGTTTTTACCTT	1141		
Db	1021	ACACTCATTTGCTGAAAGCTTAAATTAAGTACTTTCAAAA	ACGTTATCTCCATTGTTTTTACCTT	1080		
Qy	1142	CTTGAGGGGAAACGGTCTTGTGTTAAACCGCCCTGAGTTG	TCTACCCCAACCAATCTCTGTCA	1201		
Db	1081	CTTGAGGGGAAACGGTCTTGTGTTAAACCGCCCTGAGTTG	TCTACCCCAACCAATCTCTGTCA	1140		

[illegible]

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED
114076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2013)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohse, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. .2013
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:2610314P18"
/db_xref="taxon:10090"
/clone="2610314P18"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days embryo"
16. 726
/note="unnamed protein product; hypothetical RAS small
GTPases, Rab subfamily/ATP/GTP-binding site motif A
(P-loop)/Sigma-54 factor interaction domain containing
protein [InterPro|IPR003579, InterPro|IPR001687,
InterPro|IPR002078, evidence: InterPro]
putative"
/codon_start=1
/protein_id="BAC25804.1"
/db_xref="GI:26389872"
/translation="MASLDKVKVLVLGDSGVGKSLVHLIHCNQVLGNPSWTVGCSVD
IRVHDKGPEKTYVIELWDVSGVSSVSKSTRVPYNSVNGIILVHDLTNKKS
SONLYRSLVLELVNADPTGVLTNGDYDREQADNPQLPLVIGTKLDQIHETKRHEV
LITAFIAEDFNABEINLDCNTNRPSSNAVKLSRFFDKVIEKRYFFPREGNQIPG
FSDRKRFGGALKNFHCD"
1992. .1997

polyA_signal

	polyA_site	/note="putative" 2013 /note="putative"
ORIGIN		
Query Match	51.9%	Score 880; DB 3; Length 2013;
Best Local Similarity	79.1%	Pred. No. 1.7e-219;
Matches 1206; Conservative	0;	Mismatches 289; Indels 38; Gaps 12;
Qy	11	GTGCGAAGATGGGCTCCCTGGATCGGCTGAAGGTACTGTGTGGGAGACTCAGGTGTG 70
Db	8	GTGCGAAGATGGGCTCCCTGGATCGGCTGAAGGTCTGTGTGGGAGACTCAGGTGTG 67
Qy	71	GGAAATCTTGGTAGTCCATCTCTATGCCAAATCAAGTGTGGGAAATCCATCATGGA 130
Db	68	GGAAATCTTCACTCGTCCATCTCTGTGCCCAATCAAGTGTAGGAAATCCGTCATGGA 127
Qy	131	CTGTGGGCTGCTCAGTGGATGTCAGAGTTCATGATTAACAAGAGGACCCCAAGAGA 190
Db	128	CTGTGGGCTGCTCGGTGGATATCAGAGTTCATGATTAACAAGAGGACCCCTGAAGA 187
Qy	191	AGACCTACTACATAGATTAATGGATGTTGGAGCTCTGTGGGAGTCCAGCAGCGTGA 250
Db	188	AGACATATATATAGAACTATGGATGTTGGAGCTCTGTGGGAGTCCAGCAGTGTGA 247
Qy	251	AAAGCAACAGCAGTATTTCTCAACTCCGTAAATGGTATTTATTTCTACACACTTAA 310
Db	248	AAAGCAACAGCAGGTGTTCTCAACTCTGTAAACGGCATCATTTTAGTACATGACTTA 307
Qy	311	CAATAAGAGTCTCCCAAACTTGGCTGTGGTTCATTTGGAGCTCTCAACAGGATTT 370
Db	308	CAATAAGAGTCTCATCTCAAACTTATATCGTGGTCAATTTGGAAAGTTCTCAATCGGATG 367
Qy	371	TGGTGCCAACTGGAGTCTTGGTGACAAATGGGATTAATGATCAAGAACAGTTTCTGTGATA 430
Db	368	CAGTTCACCGAGTCTCTGTGACAAATGGGATTAATGATCAAGAACAGTTTCTGTGATA 427
Qy	431	ACCAATATACCACTGTTGGTAAATAGGGACTAAATCGGACCAAGATTCATGAACAAGCGCC 490
Db	428	ACCAATATCCCACTGTTGGTAAATAGGGACTAAATCGGACCAAGATTCATGAACAAGCGCC 487
Qy	491	ATGAAGTTTAACTAGGACTGCTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATT 550
Db	488	ATGAAGTTTAACTAGGACTGCTTCTCGCTGAGGACTTCAATCGTGAAGAGATTAATT 547
Qy	551	TGACATGCACAAATCCACGGTACTTAGCTGCAGCTTCTTCCAAATGCTGCAAGCTCAGTA 610
Db	548	TGATTTGCACAAACCCACGGCTCTCAGCTGAGGCTCTCAATGCTGTCAAGCTCAGTA 607
Qy	611	GGTTTTTGTATAGAGTTCATAGAGAAGATATCTTTTAAAGAGAGGTAAATCAGATTCAG 670
Db	608	GATTTTTGTATAGAGTTCATAGAGAAGATATTTTTTAAAGAGAGGTAAATCAGATTCAG 667
Qy	671	GCTTCTCTGATCGGAAAGATTTGGGCGCAGAAATTAAGAGGCTTCATTAATGACTGAA 730
Db	668	GCTTCTCTGATCGGAAAGATTTGGGCGGAGGCTTGAAGAGGCTTCAATGCTGACTGAG 727
Qy	731	TTACACTCTCTTTGGAAGAGTGACCAAGCAGTGGCAGTTTTCACAGCTCATCTTTGCT 790
Db	728	TGGCCCTCAGCC-CTGGAAGAGCAAGGAAACAGCAGGAGTGTCCAAGCTTCTCTTGTCT 786
Qy	791	GTGTTTCAATTAATTAACCATCAAGCTTTTAAACAAATCATCTTAAATATGCTTACCTTTCAG 850
Db	787	GTGTAGGACATTAATCTTCTCAGGC-TTTGGGAAACAACTCTCAATATGCTA----- 837
Qy	851	CTTTACCTTTTAATGGAATAATGAAGAGAGTGAATAACGGGAGGTCCAAATCTTTGTCC 910
Db	838	CCTCACCTCTCTTGGAAATTAAGAGAGAGCAGCGGCATCGAAGGCTCTAGCTCTGCCC 897
Qy	911	CTGTTCTCTGTGTTCTTCACTTCTGCTCCCTGTGTATAGTATATGTAAGACCTCTGTGT 970
Db	898	ATTCTCTGTGA---TCCTCACCAACCTCTGCTCTGATTAATGATTAAGAAAGCTTGTGA 954

Qy 1167 AGCCTGAGTT-GTCTACCCCAACAATCTCTGTCTTTCATTTCAAGATGCAAAATGGTGT 1225
 |||||
 Db 78 AGCCTGAGTTGGTCTACCCCAACAATCTCTGTCTTTCATTTCAAGATGCAAAATGGTGT 19
 |||||
 Qy 1226 ATTTAATGTCTCC 1239
 |||||
 Db 18 ATTTAATGTCTCC 5
 |||||

RESULT 4

AK028269
 LOCUS 2006 bp mRNA linear HTC 03-APR-2004
 DEFINITION Mus musculus 13 days embryo head cDNA, RIKEN full-length enriched library, clone:3110079H15 product:hypothetical RAS small GTPases, Rab subfamily/ATP/GTP-binding site motif A (P-loop)/Sigma-54 factor interaction domain containing protein, full insert sequence.

ACCESSION

AK028269.1 GI:26080787

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

ORGANISM Mus musculus
 Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE

4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.

AUTHORS

TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5

AUTHORS

TITLE The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 REFERENCE 6

REFERENCE

6 (bases 1 to 2006)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, K., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
 COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.jp/
 URL:http://fantom.gsc.riken.jp/.

FEATURES

source

1. 2006
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM.DB:3110079H15"
 /db_xref="taxon:10090"
 /clone="3110079H15"
 /tissue_type="head"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="13 days embryo"
 misc_feature 1. 2006
 /note="hypothetical RAS small GTPases, Rab subfamily/ATP/GTP-binding site motif A (P-loop)/Sigma-54 factor interaction domain containing protein (InterPro|IPR003579, InterPro|IPR001687, InterPro|IPR002078, evidence: InterPro)"

misc_feature

1. 2006
 /note="hypothetical RAS small GTPases, Rab subfamily/ATP/GTP-binding site motif A (P-loop)/Sigma-54 factor interaction domain containing protein (InterPro|IPR003579, InterPro|IPR001687, InterPro|IPR002078, evidence: InterPro)"

ORIGIN

Query Match 51.3%; Score 868.6; DB 3; Length 2006;
 Best Local Similarity 79.1%; Pred. No. 1.6e-216;
 Matches 1205; Conservative 0; Mismatches 279; Indels 39; Gaps 13;
 Qy 13 GGCAAGATGGCGTCCCTGGATCGGTGAAGGTACTGGTGTGGAGACTCAGGTGTGGG 72
 Db 2 GGCAAGATGGCGTCCCTGGATCGGTGAAGGTACTGGTGTGGAGACTCAGGTGTGGG 61
 Qy 73 AAATCTTCGTAGTCCATCTCTATGCAAAATCAAGTGTGGAAATCCATCATGACT 132
 Db 62 AAATCTTCATCTGTCATCTCTGTCACACATCAAGTGTAGGAATCCGTATGACT 121
 Qy 133 GTGGGCTGCTCAGTGGATGTCAGAGTTCATGATTACAAAGAGGAACCCAGAGAGAA 192
 Db 122 GTGGGCTGCTCGGTGGATATCAGAGTTCATGACTACAAAGAGGACCCCTGAAGAGA 181
 Qy 193 ACCTACTACATAGATATATGGATGTGGAGGCTCTG-TGGGCGAGTCCAGCAGCGTAA 251
 Db 182 ACATACTATATAGAACTATGGGATGTTGGAGGCTCTGCTGGCGAGTCCAGCAGTGA 241
 Qy 252 AAGCACAAGAGCAGTAGTATTTACAACTCCGTAATAGTATTATTTTCTACACGACTTAA 311
 Db 242 AAGCACAAGAGCAGGTGTTTACAACTCTGTAACCGCATCATTTTAGTACATGACTTAA 301
 Qy 312 AAATAAGAAGTCTCCCAAACTTCGTCGTGTGGTCAATGGGAAGCTCTCAACAGGATTT 371
 Db 302 AAATAAGAAGTCACTCATAAATTTATATCGTGTGTCATTGGAAGTCTCTCAATCGGAT 361
 Qy 372 GTTGCCCAACTGGAGTCTTGGTGCAAAATGGGGATTTATGATCAAGAAACAGTTTGTGATA 431
 Db 362 AGTTTCAACCGGAGTCTGGTGCAAAATGGGGATTTATGACCGGAGAACAGTTTGTGATA 421
 Qy 432 CCAATACCACTGTTGGTATAGGACTTAACCTGGACCACTTCAATCAACAGAGCCCA 491
 Db 422 CCAATCCCACTGTTGGTATAGGACTTAACCTGGACCACTTCAATCAACAGAGCCCA 481

ACCESSION	BI862084	QY	507	GACTGCTTCTCGTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGAGTGCACAAATCC	566
VERSION	BI862084.1	Db	487		
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 918)	QY	567	ACGGTACTTAGCTGCAGGTTCTTCAATGCTGTCAAGCTCAGTAGGTTTTTTTGTATAGGT	626
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .	Db	547		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)	QY	627	CATAGAGAAGATGATCTTTTAAAGAGAGGTAA--TCAGATTCAGGCTTCTCTGATCGG	684
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs-r@mail.nih.gov Tissue Procurement: DCTD/DTp cDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: InCyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12019 row: 1 column: 18 High quality sequence stop: 833. Location/Qualifiers 1. 918	Db	607	CATAGAGAAGATGATCTTTTAAAGAGAGGTAAATCAGGATTCAGGCTTCTCTGATCGG	666
FEATURES	source	QY	685	AAAGATTTGGGGCAGGACATTAAGAGGCTTCATTATGACTGATTAACATCATCCCTT	744
	/organism="Homo sapiens"	Db	667		
	/mol_type="mRNA"	QY	745	TGGAAGAGTGAGC-AAGCAGTGCCAG-TTTTTCACAGCTCATCTTGCTGTGTTTC--AAAT	800
	/db_xref="taxon:9606"	Db	727		
	/clone="IMAGE:5399969"	QY	801	ATTACATCAGAGCCTTTTAAAGATCATCTTAAATGCTACCTTACGCTTACCTT	860
	/issue_type="mammary adenocarcinoma, cell line"	Db	787	ATTACCATCAGAGTCTTTTAAAGATCATCTTAAAGATGCTACCTTACGCTTACCTT	846
	/lab_host="DH10B (phage-resistant)"	QY	861	TAATGGAAAAATGAAAGGAGTGACAAATACGGGAGTCCAAACTTTGTCCCTGTTCTCTG	920
	/clone_lib="NIH MGC 87"	Db	847	ATGGCACAACCTGAACGGGAAGTGACATACGGAAGGTCCAAATTTGTCCCTGTTCTCTC	906
	/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."	QY	921	TGT 923	
		Db	907	TGT 909	
ORIGIN					
Query Match	46.7%; Score 791.8; DB 4; Length 918;				
Best Local Similarity	95.2%; Pred. No. 1.9e-196;				
Matches	860; Conservative 0; Mismatches 37; Indels 6; Gaps 4;				
QY	27	CCTGGATCGGTCAAGTACTGCTGTGGAGACTCAGGTGTGGAAATCTTCGTAGT	86		
Db	7	CCTGGATCGGTGAAGTACTGCTGTGGAGACTCAGGTGTGGAAATCTTCGTAGT	66		
QY	87	CCATCTCCTATGCAAAATCAAGTGTGGAAATCCATCATGACTGTGGCTGCTCAGT	146		
Db	67	CCATCTCCTATGCAAAATCAAGTGTGGAAATCCATCATGACTGTGGCTGCTCAGT	126		
QY	147	GGATGTCAGATTCATGATTACAAAGAGAACCCAGAGAGAACCTTACATAGTA	206		
Db	127	GGATGTCAGATTCATGATTACAAAGAGAACCCAGAGAGAACCTTACATAGTA	186		
QY	207	ATTATGCGATTTGGAGGCTCTCTGGGAGTGCAGCGCTCAAAAGCACAGAGCAGT	266		
Db	187	ATTATGCGATTTGGAGGCTCTCTGGGAGTGCAGCGCTCAAAAGCACAGAGCAGT	246		
QY	267	ATTCTCAACTCGGTAATGGTATTATTTTCGTACAGCTTAAACAAATGAAGTCTTC	326		
Db	247	ATTCTCAACTCGGTAATGGTATTATTTTCGTACAGCTTAAACAAATGAAGTCTTC	306		
QY	327	CAAAACTCTGCTGCTGTTGTCATGAGAGTCTCAACAGGATTTGGTGCACCTGGAGT	386		
Db	307	CAAAACTCTGCTGCTGTTGTCATGAGAGTCTCAACAGGATTTGGTGCACCTGGAGT	366		
QY	387	CTTGGTGACAAATGGGATTTATCATCAAGAACAGTTTGTGTGATAACCAATACCTGTT	446		
Db	367	CTTGGTGACAAATGGGATTTATCATCAAGAACAGTTTGTGTGATAACCAATACCTGTT	426		
QY	447	GGTAATAGGAGCTAAACTGGACAGATTCATGAAACAAAGCGCCATGAAGTTTAACTAG	506		
Db	427	GGTAATAGGAGCTAAACTGGACAGATTCATGAAACAAAGCGCCATGAAGTTTAACTAG	486		
LOCUS	AK011280				
DEFINITION	Mus musculus 10 days embryo whole body cDNA, RIKEN full-length enriched library, clone:261000D21 product:hypothetical RAS small GTPases, Rab subfamily/ATP/GTP-binding site motif A (P-loop)/Sigma-54 factor interaction domain containing protein, full insert sequence.				
ACCESSION	AK011280				
VERSION	AK011280.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20493374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Fogawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				

JOURNAL MEDLINE PUBMED	Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861
REFERENCE AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Nature 409, 685-690 (2001)
TITLE JOURNAL REFERENCE AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 1961) Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukushima, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE JOURNAL	Direct Submission Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGATCTCGAGTAAATAAATTAATCCCCCCCCC 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATCTCGAGTAAATAAATTAATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SacII. Cloning sites, 5' end: XhoI; 3' end: SacII. Host: SOLR.
FEATURES source	Location/Qualifiers 1..1961 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:2610002D21" /db_xref="taxon:10090" /clone="261002D21" /issue_type="whole body" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="10 days embryo" 17..406 /notes="unnamed protein product; hypothetical RAS small GTPases, Rab subfamily/ATP/GTP-binding site motif A (P-loop)/Sigma-54 factor interaction domain containing protein (InterPro IPR003579, InterPro IPR001687, InterPro IPR02078, evidence: InterPro) putative" /codon_start=1 /protein_id="BAB27514.1" /db_xref="GI:12847299"
CDS	/translation="MASLDRVKVLVLDGVGKSLVHLLCHNQLVGNPNMTWGCSVD IRVHDYKEGTPEEKTYIRLEMDVGSGSVGSVKSVAFTVPFNVSNGILVHDLTNKES SONLYRWLSLEVNDAVPTGLVIINGD"

```
QY 971 AAATATGAGATGTTGTCAAAATGATCAGTAAATGACCAATGACAGTGTACTGACAGAA 1030
Db 903 CGACATAAGGCTGTGCAAAACCATGAGTAAGAACTGAGAGACA-AGGCTGCACGAA 961
QY 1031 AATTTACTCTGTGCTAGAACTGAGGGTTTTTATAGGCTCTGTAATTTTCCACACTCAT 1090
Db 962 GGCTGACTCTTACTTGGAAACGCGGG-TTTTACAGGCTGCGATTTTCCACACTTGCT 1020
QY 1091 GCTGAAGCTTAATTAAGTACTTCAAAAGGATCTCCATTTGTTTACCTTCTTGAGGG 1150
Db 1021 GCTGAAGGCTTAATTAAGTACTTCAAAACCATCTGTGTGTTTAA-CTTCTGAGGG 1079
QY 1151 AAGGCTTTGTTAAACGAGCTGAGTTGTCTACCCCAAAACAATCTCTGTCAATTTTCAAAG 1210
Db 1080 AAGGTCATGTTAATCAGGCTGAGTTATCCACCCAAACCATCTCTGCAATTTTAAAG 1139
QY 1211 ATGCAAAATGGTGTATTTAATTTGTTCTCCACCAATGTCACACACAGGAATGCCATAAT 1270
Db 1140 AAGCCAAATGGTGTATTTCAATTTTCTCCATCTCATTAACACAGGGATGCCAAGAT 1199
QY 1271 AGCAACCTTGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1327
Db 1200 AGCAACCC--CATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1257
QY 1328 GGACTAATAGCCAGAGTAAATATAATACAAATTAATATAATACATAGAACAGCAATAC 1387
Db 1258 GGACTAATAGCTGGAGTAAATAGAAATACAAATTAATATAATACATGGCAACAG----AC 1313
QY 1388 CAGAAAAAAGAAATTCGTGTAAATGATGTGAAAAATTCAGACCTCCCTCACTCTTAAGG 1447
Db 1314 CAGAGAGAGGGT-----CTGTGAGACTTAACAGCTCTCTGCTC-CAAGG 1360
QY 1448 TTGCTGCTATATACAGTCTAGGTTTCTGTTTGGAAATAGGTAGGTAAATCTAAGACC 1507
Db 1361 TTGAGCTGTGGAAGGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1420
QY 1508 TGCACAAGGCGCTGAGAGACATT 1531
Db 1421 TGTAAGGCCAGTAAGACAGATT 1444
```

```
RESULT 9
BQ215951
LOCUS BQ215951 815 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7558711 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6045318
5', mRNA sequence.
ACCESSION BQ215951
VERSION BQ215951.1 GI:20397351
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M13289 row: f column: 07
High quality sequence stop: 614.
Location/Qualifiers
1..815
/organism="Homo sapiens"
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="IMAGE:6045318"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/Note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match 44.3%; Score 750.6; DB 5; Length 815;
Best Local Similarity 98.4%; Pred. No. 1.2e-185;
Matches 779; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

QY 51 GTTGGAGACTCAGGTGTTGGGAAATCTTCGTTAGTCCATCTCTATGCCAAATCAAGT 110
Db 1 GTTGGAGACTCAGGTGTTGGGAAATCTTCGTTAGTCCATCTCTATGCCAAATCAAGT 60
QY 111 GCTGGAAATCCATCATGGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 170
Db 61 GCTGGAAATCCATCATGGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 171 AGAAGGAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 230
Db 121 AGAAGGAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 231 GGGAGAGTCCAGCAGCGTGAAGAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 290
Db 181 GGGAGAGTCCAGCAGCGTGAAGAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 291 TATTTTCTGACAGGCTTAACAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 350
Db 241 TATTTTCTGACAGGCTTAACAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 351 GGAAGCTCTCAACAGGAGTTTGGTGCAACTGGAGTCTTGGTGCAAAATGGGAGATTATGA 410
Db 301 GGAAGCTCTCAACAGGAGTTTGGTGCAACTGGAGTCTTGGTGCAAAATGGGAGATTATGA 360
QY 411 TCAAGAAAGTGTGCTGATAACCAATACCACTGTTGGTAAATAGGAGCTAAACTGGACCA 470
Db 361 TCAAGAAAGTGTGCTGATAACCAATACCACTGTTGGTAAATAGGAGCTAAACTGGACCA 420
QY 471 GATTATGAAACAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 530
Db 421 GATTATGAAACAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 480
QY 531 CAATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 590
Db 481 CAATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 540
QY 591 CAATGCTGCAAGCTCAGTGTGTTTTCATAGGTCATAGAGAGAGAGAGAGAGAGAGAG 650
Db 541 CAATGCTGCAAGCTCAGTGTGTTTTCATAGGTCATAGAGAGAGAGAGAGAGAGAGAG 600
QY 651 AGAAGGTAATCAGATTCAGGCTTTCCTGATCGGAAAGAGATTGGGCGAGGAGAGATTAA 710
Db 601 AGAAGGTAATCAGATTCAGGCTTTCCTGATCGGAAAGAGATTGGGCGAGGAGAGATTAA 660
QY 711 GAGCCTTCATATGACTGAATTTACACTCATCTTTTGGAGAGAGAGAGAGAGAGAGAGAG 770
Db 661 GAGCCTTCATATGACTGAATTTACACTCATCTTTTGGAGAGAGAGAGAGAGAGAGAGAG 720
QY 771 TTTTCAAGCTCATCTTGCT-GTGTTCATATTATACCATCAC--AGCCTTTTAAACAAAT 837
Db 721 TTTTCAAGCTCATCTTGCTGGGTTCAATTATTATACCATCAC--AGCCTTTTAAACAAAT 780
QY 837 CATCTTAAATG 839
Db 781 CATCTTAAATG 792
```

RESULT 10
 CK456385/c
 LOCUS
 DEFINITION 875 bp mRNA linear EST 12-JAN-2004
 CK456385
 ACCESSION 917058 MARC 4PIG Sus scrofa cDNA 3', mRNA sequence.
 VERSION CK456385.1 GI:40803599
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 875)
 AUTHORS Smith,T.P.L., Preking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
 Nonneman,D.O., Wray,J.E. and Keele,J.W.
 TITLE Porcine EST collection using a normalized library constructed from
 embryos representing early developmental stages
 JOURNAL Unpublished (2003)
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim.alt option. Vector identified with
 cross match v0.990329.
 Plate: TWM8028 row: B column: 12
 Seq primer: TAGAAGGCACAGTCGAGG.
 Location/Qualifiers
 1..875
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="WARC 4PIG"
 /note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
 Library made with combined RNA from day-10, day-13,
 day-15, day-25, and day-30 whole embryos."

FEATURES

source

Query Match 44.3%; Score 750; DB 7; Length 875;
 Best Local Similarity 91.9%; Pred. No. 1.8e-185;
 Matches 803; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
 QY 101 AAAATCAAGTCTGGGAAATCCATCATGAGTCTGGGCTGCTCAGTGGATGTCAGAGTTC 160
 DB 875 AAATCAAGTCTGGGAAATCCATCATGAGTCTGGGCTGCTCAGTGGATGTCAGAGTTC 816
 QY 161 ATGATTACAAAGAGGAACCCCAAGAGAGACCTACTACATAGAAATTATGGGATGTTG 220
 DB 815 ATGACTACAAAGAGGAACCCCAAGAGAGACCTACTATATAGAAATTATGGGATGTTG 756
 QY 221 GAGGCTCTGGGAGTGCAGAGCGTGAAGAGCAAGAGCGATGTTCTACAACTCG 280
 DB 755 GAGGCTCGTGGGAGTGCAGAGCGTGAAGAGCAAGAGCGATGTTCTACAACTCTG 696
 QY 281 TAAATGTTATTTTCGTACAGACTTAACAATAGAGTCTCCCAAACTTGGTC 340
 DB 695 TAAATGTTATTTTGTATGATGACTTAACAATAGAGTCTCCCAAACTTGGTATC 636
 QY 341 GTTGGTCATTGGAAGCTCTCAACAGGGATTTGGTGCCAACTGGAGTCTTGGTACAAATG 400
 DB 635 GTTGGTCATTGGAAGCTCTCAACAGGGATTTGGTGCCAACTGGAGTCTTGGTACAAATG 576
 QY 401 GGGATTATGATCAAGAACAGTGTCTGATACCAATACCACTGTTGGTAAATAGGACTA 460
 DB 575 GGGATTATGACCGGAAACAGTGTCTGATACCAATACCACTGTTGGTAAATAGGACTA 516
 QY 461 AACTGGACCAAGATTCATGAACAAAGCGCCATGAAGTTTAACTAGGACTGCTTCTCGG 520
 DB 515 AACTGGATCAGATTCATGAACAAAGCGCCATGAAGTTTAACTAGGACTGCTTCTCGG 456

QY 521 CTGAGGATTTCAATCAGAGAGAAATTAATTTGGACTGCACAAATCCAGGTACTTACTG 580
 DB 455 CTGAGGATTTCAATCAGAGAGAGATTAATTTGGATTGACAAATCCAGGTACTTACTG 396
 QY 581 CAGGTTCTTCCAAATGCTGCAAGCTCAGTAGGTTTTTTTGAATAAGGTCATAGAGAAGAT 640
 DB 395 CAGGTTCTTCCAAATGCTGTCNAGCTCAGTAGGTTTTTTTGAATAAGGTCATAGAGAAGAT 336
 QY 641 ACTTTTAAAGAGAGGTAATCAGATCCAGGCTTCTGATCGGAAAGATTTGGGGCAG 700
 DB 335 ACTTTTAAAGAGAGGTAATCAGATCCAGGCTTCTGATCGGAAAGATTTGGGGCAG 276
 QY 701 GACATTAAGAGAGCTTCAATATGACTGAATACATCATCCCTTGAAGAGTGAGCAAG 760
 DB 275 GAACATTAAGAGAGCTTACATTAATGACTGAATACATCATCCCTTGAAGAGTGAGCAAT 216
 QY 761 CAGTGGCAGTTTTTTCAGC-TCATCTTGTCTGTTTCAATTAATACCATCACACCTTTT 819
 DB 215 TGGTGACAGTTTTTTCAGAGGTTCTCTCTGTTTGAATATTCCTGTCCAGCCTTTT 156
 QY 820 AACAAATCATCTTAAATGCTACCTTTCAGCCTTACCTTTTAATGAAAAAATGAAAGGA 879
 DB 155 AGCAAAATCATCTTAAATGCTACCTTTCAGCCTTTCAGCCTTTCAGCCTTTCAGCCTTTC 96
 QY 880 AGTGCACATACGAGAGTCCAAACTTGTCCCTGTTCTCTGTTCTCTGTTCTCTGTTCTCTGTC 939
 DB 95 AGTGCACAGCGTGGGAGTCCAAACTTGTCCCTGTTCTCTGTTCTCTGTTCTCTGTTCTCTGTC 36
 QY 940 CTTGTGTATAGATTATGTAAAGCCTTGTGTAA 973
 DB 35 CCTTTTATAGATTATATAAAGCCTTGGGAAA 2

RESULT 11

BQ889351

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..950

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6179062"

/sex="male"

/tissue_type="dorsal root ganglia"

/dev_stage="adult, 36 yr"

/lab_host="DH10B"

/clone_lib="Lupski_dorsal root ganglion"

/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:

BQ889351

950 bp mRNA linear EST 16-AUG-2002

AGENCOURT 8113408 Lupski_dorsal_root_ganglion Homo sapiens cDNA

clone IMAGE:6179062 5', mRNA sequence.

BQ889351

BQ889351.1 GI:22281365

EST.

Homo sapiens (human)

Homo sapiens

Homo sapiens

Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Dr. James R. Lupski

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13559 row: j column: 23

High quality sequence start: 19

High quality sequence stop: 639.

Location/Qualifiers

1..950

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6179062"

/sex="male"

/tissue_type="dorsal root ganglia"

/dev_stage="adult, 36 yr"

/lab_host="DH10B"

/clone_lib="Lupski_dorsal root ganglion"

/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:

NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGTCCG-3' and
5'-GACTAGTCTTAGTCGCGAGCGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN		Query Match		44.1%; Score 748.2; DB 5; Length 950;		Best Local Similarity		97.3%; Pred. No. 5.5e-185;		Matches 793; Conservative 0; Mismatches 18; Indels 4; Gaps 3;	
QY	876	AGGAAGTGACATACGCGGAGGTCCAAACTTTGTCCTCTCTCTGTCCTTACCTTTC	935								
DB	24	ATGAAGTGACATACGCGGAGGTCCAAACTTTGTCCTCTCTCTGTCCTTACCTTTC	83								
QY	936	TGTCCCTGTGTAGATATTATGTAAGCCCTTGTAATATGATGTTGTCAAAATGAT	995								
DB	84	TGTCCCTGTGTAGATATTATGTAAGCCCTTGTAATATGATGTTGTCAAAATGAT	143								
QY	996	GCAGTAAATGAGCAATACAGAGTACTCTGCAGAGAAAATTTACTCTTGCCTAGAACTGGAG	1055								
DB	144	GCAGTAAATGAGCAATACAGAGTACTCTGCAGAG-NAATTACTCTTGCCTAGAACTGGAG	202								
QY	1056	GGTTTTATGGGTCTGTAAATTTTCCCACTCATCTGCTGAAGCTTAATTAAGTACTTCA	1115								
DB	203	GGTTTTATGGGTCTGTAAATTTTCCCACTCATCTGCTGAAGCTTAATTAAGTACTTCA	262								
QY	1116	AAACGTAATCTCATTTGTTTACCTCTTCTGAGGGAGCGCTTGTAAACCGCCCTGAG	1175								
DB	263	AAACGTAATCTCATTTGTTTACCTCTTCTGAGGGAGCGCTTGTAAACCGCCCTGAG	322								
QY	1176	TTGCTACCCCAACAACTCTCTCATTTTCAAGATGCAAAATGCTGTTATTAATTGT	1235								
DB	323	TTGCTACCCCAACAACTCTCTCATTTTCAAGATGCAAAATGCTGTTATTAATTGT	382								
QY	1236	CTCCACATTTGTACACACAGGAATGCTTAATTAAGCAACCTTGTCTCCCTCTCTCT	1295								
DB	383	CTCCACATTTGTACACACAGGAATGCTTAATTAAGCAACCTTGTCTCCCTCTCTCTCT	442								
QY	1296	CCTTTCGAAATGGCTCAGTGACCTGGAAGGCGGACTTAATAGCCAGTTAATATAAAT	1355								
DB	443	CCTTTCGAAATGGCTCAGTGACCTGGAAGGCGGACTTAATAGCCAGTTAATATAAAT	502								
QY	1356	ACAAATTAATAATACATAGAGAACAGCAATACCAGAAAAAGAAATCTGTTAAATGAT	1415								
DB	503	ACAAATTAATAATACATAGAGAACAGCAATACCAGAAAAAGAAATCTGTTAAATGAT	562								
QY	1416	GTGAAAAATTGACAGCTCCCTCACTCTTAAGGTGTGCTATATACAGTCTAGGTTTCT	1475								
DB	563	GTGAAAAATTGACAGCTCCCTCACTCTTAAGGTGTGCTATATACAGTCTAGGTTTCT	622								
QY	1476	GTTTGCAATAGGTAGGTGTAATCTTAAGCCTGCAAGGGCAGTGAGAGCATTTTACA	1535								
DB	623	GTTTGCAATAGGTAGGTGTAATCTTAAGCCTGCAAGGGCAGTGAGAGCATTTTACA	682								
QY	1536	G-CCTCTCTCTATTGTTTTTTTAAAGGAAAAGTCAACTCC--TGAATATGTCCTTAGCT	1592								
DB	683	GCCCTCTCTATTGTTTTTTTAAAGGAAAAGTCAACTCCCTGGAATGTCCTTAGCT	742								
QY	1593	ATAATCAGAAAACTAAGATATTATTCTGTGTCAACAAATGATATTATGAGAGAGATAA	1652								
DB	743	ATAATCAGAAAACTAAGATATTATTCTGTGTCAACAAATGATATTATGAGAGAGATAA	802								
QY	1653	AATAAGTTCACAGCAACACAAAACATGATTAAT	1687								
DB	803	AATAAGTTCACAGCAACACATTTACATGAATTAAT	837								

CK455329
LOCUS 915906 MARC 4PIG Sus scrofa cDNA 5', mRNA linear EST 12-JAN-2004
DEFINITION CK455329
ACCESSION CK455329
VERSION 1 GI:40802543
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 855)
AUTHORS Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross match v0.990329.
Plate: TMM8028 row: B column: 12
Seq primer: GTAATACGACTCACTATAGG.
Location/Qualifiers
1..855
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4PIG"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN

Query Match		44.1%; Score 747; DB 7; Length 855;		Best Local Similarity		92.9%; Pred. No. 1.1e-184;		Matches 794; Conservative 0; Mismatches 60; Indels 1; Gaps 1;	
QY	24	GTCCCTGGATCGGTGAAGGTACTGGTGTGGAGACTCAGGTGTGGGAATCTTCGT	83						
DB	1	GTCCCTGGATCGGTGAAGGTACTGGTGTGGAGACTCAGGTGTGGGAATCTTCCT	60						
QY	84	AGTCCATCTCTATGCGCAAAATCAAGTCTGGGAATCCATCATGACTGTGGCTGCTC	143						
DB	61	AGTTCATCTTCTATGCGCAAAATCAAGTCTGGGAATCCATCATGACTGTGGCTGCTC	120						
QY	144	AGTGGATGTCAGAGTTTCATGATTACAAAGAGAAACCCAGAGAGAACCTTACAT	203						
DB	121	AGTAGATGTCAGAGTTTCATGACTACAAAGAGAAACCCAGAGAGAACCTTACTAT	180						
QY	204	AGAAATATGGGATTTGGAGGCTCTGTGGCAGTGCAGAGCGTCAAAAGCAGAGGC	263						
DB	181	AGAAATATGGGATTTGGAGGCTCGGTGGGAGTGCAGAGTGTGAAAGCAGAGGC	240						
QY	264	AGTATTCTACAACTCGTAAATGGTATTATTTTGTACACGACTTAACTAAAGAGTC	323						
DB	241	AGTATTCTACAACTCGTAAATGGTATTATTTAGTACATGACTTAACTAAAGAGTC	300						
QY	324	CTCCCAAACTTTCGCTCGTTGTCATTGGAAGCTCTCAACAGGGATTTGGTGCAC	383						
DB	301	ATCCCAAAATTTGTATCGTTGTCATTGGAAGCTCTCAACAGGGATCTGGTGCAC	360						
QY	384	AGTCTTGTGACAAATGGGGATTATGATCAAGAACAGTTTGTGATAACCAATAC	443						
DB	361	AGTCTTGTGACAAATGGGGATTATGATCAAGAACAGTTTGTGATAACCAATAC	420						
QY	444	GTTGGTAATAGGGAATTAACCTGGACCGATTCATGAAACAAAGCGCCATGAAGT	503						

```
Db 421 GTTGGTAATAGGAGCTAACTGGATCAGATTCATGAACCAAGCGCCATGAAGTTTAAAC 480
Qy 504 TAGGACTGCTTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGCACAAA 563
Db 481 TAGGACTGCTTTCTCGCTGAGGATTTCAATGCAGAGAGATTAATTTGGATTGCACAAA 540
Qy 564 TCACGGTACTTAGCTGAGGCTTCTTCATAGCTGTCAGCTCAGTCAAGTCTTCTGATTA 623
Db 541 TCACGGTACTTAGCTGAGGCTTCTTCATAGCTGTCAGCTCAGTCAAGTCTTCTGATTA 600
Qy 624 GGTATAGAGAGAGATTAATTTTAAAGAGAGGTAATCAGATTCAGGCTTCTCGATCG 683
Db 601 GGTATAGAGAGAGATTAATTTTAAAGAGAGGTAATCAGATTCAGGCTTCTCGATCG 660
Qy 684 GAAAGATTTTGGGCGAGCAATTAAGAGAGCTTCAATATGACTGAATTAACATCATCT 743
Db 661 AAAAGGTTTGGAGGAGGACATTAAGAGAGCTTCAATATGACTGAATTAACATCATCT 720
Qy 744 TTGGAAGAGTGAAGAGAGGAGGAGGCTTTTTCACAGC-TCATCTTGCTGTGTTCAATAT 802
Db 721 TTGGAAGAGGAGCAATTTGCTGACAGTCTTTTTCACAGGCTTCTCGCTGTTTCGATTAT 780
Qy 803 TACCATCAGAGCTTTTAAACAAATCATCTTAAATGCTACCTTCAGGCTTACCTTTTA 862
Db 781 TCCTGCTCCAGGCTTTTAGCAAAATCATCTTAAATGCTACCTTCAGGCTTACCTTTTA 840
Qy 863 ATGGAATAATCAAG 877
Db 841 ATGGAATAATCAAG 855
```

```
RESULT 13
LOCUS CN361073 720 bp mRNA linear EST 16-MAY-2004
DEFINITION 328775509 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN361073
VERSION CN361073.1 GI:47361007
SOURCE EST.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 720 Std Error: 0.00.
```

```
FEATURES
source
1..720
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone_lib="GRN_ES"
/note="oligo dt primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"
```

ORIGIN

Query Match 42.4%; Score 718.4; DB 7; Length 720;
Best Local Similarity 99.9%; Pred. No. 3.4e-177;

```
Matches 719; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 49 GTGTGGGAGACTCAGGTGTGGGAAATCTTCGTTAGTCCATCTCTATGCGCAAAATCAA 108
Db 1 GGGTTGGGAGACTCAGGTGTGGGAAATCTTCGTTAGTCCATCTCTATGCGCAAAATCAA 60
Qy 109 GTGTGGGAAATCCATCATGAGCTGTGGGCTGTGCTGAGTGTGAGGATTCAGAGTTCAATGATAC 168
Db 61 GTGTGGGAAATCCATCATGAGCTGTGGGCTGTGCTGAGTGTGAGGATTCAGAGTTCAATGATAC 120
Qy 169 AAAGAAGGAAACCCAGAGAGAGAGAGCTTACTACATAGAAATTTATGGGATGTTGGAGGCTCT 228
Db 121 AAAGAAGGAAACCCAGAGAGAGAGAGCTTACTACATAGAAATTTATGGGATGTTGGAGGCTCT 180
Qy 229 GTGGGCTGAGTGTGCTGATTAACCAATACCACTGTGTTGTAATAGGAGCTAAACTGGAC 288
Db 181 GTGGGCTGAGTGTGCTGATTAACCAATACCACTGTGTTGTAATAGGAGCTAAACTGGAC 240
Qy 289 ATTATTTTCTGACAGGCTTAAACAAATAAGAGGCTCTCCCAAACTTTCGCTGTTGGTCA 348
Db 241 ATTATTTTCTGACAGGCTTAAACAAATAAGAGGCTCTCCCAAACTTTCGCTGTTGGTCA 300
Qy 349 TTGGAAGCTCTCAACAGGCTTGGTGCCTGCTGAGTCTTGGTGAACAATGGGATTAAT 408
Db 301 TTGGAAGCTCTCAACAGGCTTGGTGCCTGCTGAGTCTTGGTGAACAATGGGATTAAT 360
Qy 409 GATCAAGAACAGTTTGTGATTAACCAATACCACTGTGTTGTAATAGGAGCTAAACTGGAC 468
Db 361 GATCAAGAACAGTTTGTGATTAACCAATACCACTGTGTTGTAATAGGAGCTAAACTGGAC 420
Qy 469 CAGATTTCATGAACAAAGCGCATGAAGTTTAACTAGGAGCTGCTTTCCTGCTGAGGAT 528
Db 421 CAGATTTCATGAACAAAGCGCATGAAGTTTAACTAGGAGCTGCTTTCCTGCTGAGGAT 480
Qy 529 TTCAATCCAGAGAAATTAATTTGCTGCTGCACTGCACTGCACTGCTGCTGCTGCTGCTGCT 588
Db 481 TTCAATCCAGAGAAATTAATTTGCTGCTGCACTGCACTGCACTGCTGCTGCTGCTGCTGCT 540
Qy 589 TCCAATGCTGCAAGCTCAGTAGGCTTTTGTATAGGCTCATAGAGAGAGATACCTTTTGA 648
Db 541 TCCAATGCTGCAAGCTCAGTAGGCTTTTGTATAGGCTCATAGAGAGAGATACCTTTTGA 600
Qy 649 AGAGAAGGTAATCAGATTCAGGCTTTCCTGATCGGAAAGATTTGGGGAGGAGCACTTA 708
Db 601 AGAGAAGGTAATCAGATTCAGGCTTTCCTGATCGGAAAGATTTGGGGAGGAGCACTTA 660
Qy 709 AAGAGCCTTCATATGACTGAATTTACATCTCTTTTGGAGAGCTGAGCAAGCAGTGGCA 768
Db 661 AAGAGCCTTCATATGACTGAATTTACATCTCTTTTGGAGAGCTGAGCAAGCAGTGGCA 720
```

RESULT 14

LOCUS BX099226 718 bp mRNA linear EST 06-FEB-2003
DEFINITION BX099226 Soares_fetal_liver_splice_infls_sl Homo sapiens cDNA clone
IMAGE9981061023; IMAGE:435149, mRNA sequence.

ACCESSION BX099226
VERSION BX099226.1 GI:27843984
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 718)
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
Human Unigeneset - RZPD3
Contact: Ina Rolfs
Unpublished (2003)
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE9981061023.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;

Human Unigeneset - RZPD3 (RZPDLIB No. 972)
<http://www.rzpd.de/Cloncards/cgi-bin/showlib.pl.cgi?responseLibNo=372> Contact: Ina Rolfe
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 101
 Fax: +49 30 32639 111
www.rzpd.de
 This clone is available royalty-free from RZPD;
 contact RZPD (clone@rzpd.de) for further information. Seq primer:
 M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.
 Location/Qualifiers
 1. .718
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE9981061023 ; IMAGE:435149"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="PH10B (ampicillin resistant)"
 /clone_lib="SOARES fetal liver spleen INFLS S1"
 /notes="Organ: Liver and Spleen; Vector: p7T73D (Pharmacia)
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
 This is a subtracted version of the original Soares fetal
 liver spleen INFLS library. 1st strand cDNA was primed
 with a Pac I - oligo (dT) primer [5'
 AAGTGAAGATTAATTAAGATCTTTTTTTTTTTTTTTT 3']
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified p7T73 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Ronaldo."

FEATURES

source

ORIGIN

Query Match	42.4%;	Score 718;	DB 5;	Length 718;
Best Local Similarity	100.0%;	Prod. No. 4.3e-177;		
Matches 718;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	46	CTGCTGTTGGAGACTCAGGTGTTGGAAATCTTCCTGTTAGTCCATCTCCTATGCAAAAT	105	
DB	1	CTGCTGTTGGAGACTCAGGTGTTGGAAATCTTCCTGTTAGTCCATCTCCTATGCAAAAT	60	
QY	106	CAAGTCTCGGAAATCCATCATGACTGTGGGCTGCTCAGTGGATGTGAGATTCATGAT	165	
DB	61	CAAGTCTCGGAAATCCATCATGACTGTGGGCTGCTCAGTGGATGTGAGATTCATGAT	120	
QY	166	TACAAAGAGAACCCAGAGAGAACCTTACATAGATAATATGGGATGTTGGAGGC	225	
DB	121	TACAAAGAGAGAACCCAGAGAGAACCTTACATAGATAATATGGGATGTTGGAGGC	180	
QY	226	TCTGTGGGAGTGCACGACGGTCAAAAGACACAAGACGAGTATTTCTACACTCCGTAAT	285	
DB	181	TCTGTGGGAGTGCACGACGGTCAAAAGACACAAGACGAGTATTTCTACACTCCGTAAT	240	
QY	286	GGTATTATTTTCGTACACGACTTAACAAATAAGAAAGTCTCCCAAAACTTTGCGTCGTGG	345	
DB	241	GGTATTATTTTCGTACACGACTTAACAAATAAGAAAGTCTCCCAAAACTTTGCGTCGTGG	300	
QY	346	TCATTGGAAGCTCTCAACAGGGATTTGGTGCCAACTGGAGTCTTTGGTGA CAAAATGGGAT	405	
DB	301	TCATTGGAAGCTCTCAACAGGGATTTGGTGCCAACTGGAGTCTTTGGTGA CAAAATGGGAT	360	
QY	406	TATGATCAAGAACAGTTTGTGTGATAACCAATACACATGTTGGTAAATAGGGACTAAACTG	465	
DB	361	TATGATCAAGAACAGTTTGTGTGATAACCAATACACATGTTGGTAAATAGGGACTAAACTG	420	
QY	466	GACCAGATTTCATGAAACAAGCGGCATGAAAGTTTTTAATAGGACTGCTTTCTGGCTGAG	525	
DB	421	GACCAGATTTCATGAAACAAGCGGCATGAAAGTTTTTAATAGGACTGCTTTCTGGCTGAG	480	
QY	526	GATTTCAATCCAGAAGAAATTTAATTTGGACTGCA CAAATCCAGGTACTTTAGTGCAGGT	585	
DB	481	GATTTCAATCCAGAAGAAATTTAATTTGGACTGCA CAAATCCAGGTACTTTAGTGCAGGT	540	

Qy	586	TCTTCCAATGCTGTCAAGCTCAGTAGGTTTTTTGATAAGGTCATAGAGAAGAGATACCTTT	645
Db	541	TCTTCCAATGCTGTCAAGCTCAGTAGGTTTTTTGATAAGGTCATAGAGAAGAGATACCTTT	600
Qy	646	TTAAGAGAAGGTAATCAGATTCACAGCTTCCATGATCGGAAAGAGATTTGGGGCAGGAACA	705
Db	601	TTAAGAGAAGGTAATCAGATTCACAGCTTCCATGATCGGAAAGAGATTTGGGGCAGGAACA	660
Qy	706	TTAAGAGAGCTTCATTATGACTGAATATACACTCATCTCTTTGGAAGAGTGTAGCAAGCAG	763
Db	661	TTAAGAGAGCTTCATTATGACTGAATATACACTCATCTCTTTGGAAGAGTGTAGCAAGCAG	718
RESULT 15			
BM451532			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
Qy	63	AGGTGTTGGGAATCTTCGTTAGTCATCTCCATGCGCAAAATCAAGTGTGGGAATCC	122
Db	182	AGGTGTTGGGAATCTTCGTTAGTCATCTCCATGCGCAAAATCAAGTGTGGGAATCC	241
Qy	123	ATCATGCACTGTGGCTGTCTCAGTGGATGTACAGATTCATGATTACAAAGAGAACCCC	182
Db	242	ATCATGCACTGTGGCTGTCTCAGTGGATGTACAGATTCATGATTACAAAGAGAACCCC	301
Qy	183	AGAAGAGAAGACCTACTACATAGATTTATGGATGTTGGAGGCTCTGTGGCAGTGCAG	242
Db	302	AGAAGAGAAGACCTACTACATAGATTTATGGATGTTGGAGGCTCTGTGGCAGTGCAG	361
Qy	243	CAGCGTGAAAGACACAGAGCAGTATTTCTTCAACATCCGTAATGGTATTTATTTTCGTACA	302

362	Db		CAGCGTGA	AAGCA	CAAGAG	CAGTATCT	CAAC	TTCG	TAAATG	TATATTTTCG	TACA	421					
303	Qy		CGACTTAA	CAAAAT	AAGAAG	CCTCC	CAAA	CTTG	CGTGGT	CTCATTTG	GGAAGCTCT	CA	362				
422	Db		CGACTTAA	CAAAAT	AAGAAG	TCTCC	CAAA	CTTG	CGTGGT	CTCATTTG	GGAAGCTCT	CA	481				
363	Qy		CAGGGATTT	TGGTGC	CAACTG	GAGTCTT	GGTG	CAAA	TGGG	GAATAT	TGATCA	AAGAACAG	422				
482	Db		CAGGGATTT	TGGTGC	CAACTG	GAGTCTT	GGTG	CAAA	TGGG	GAATAT	TGATCA	AAGAACAG	541				
423	Qy		TGCTGATA	CAAAAT	ACA	CTGTT	TGGTAA	TAGG	GA	CTAAA	CTGG	ACAGATTC	ATGA	AAAC	482		
542	Db		TGCTGATA	CAAAAT	ACA	CTGTT	TGGTAA	TAGG	GA	CTAAA	CTGG	ACAGATTC	ATGA	AAAC	601		
483	Qy		AAAGCG	CGCATGA	AGTTTT	TAACT	AGG	AC	TG	CTGCT	GAGG	ATTTCA	ATCC	AGA	542		
602	Db		AAAGCG	CGCATGA	AGTTTT	TAACT	AGG	AC	TG	CTGCT	GAGG	ATTTCA	ATCC	AGA	661		
543	Qy		AAATTAATTT	TGGAC	TGC	CAAA	ATCC	ACG	GTA	CTTAG	CTCG	AGGTTCTT	CCAA	TGCTG	CA	602	
662	Db		AAATTAATTT	TGGAC	TGC	CAAA	ATCC	ACG	GTA	CTTAG	CTCG	AGGTTCTT	CCAA	TGCTG	CA	721	
603	Qy		GCTCAG	TAGGTTTT	TGATA	AGGTC	CATAG	AGA	AGAT	TACTTTTT	TAA	GAGAAG	AGGTA	ATCA	662		
722	Db		GCTCAG	TAGGTTTT	TGATA	AGGTC	CATAG	AGA	AGAT	TACTTTTT	TAA	GAGAAG	AGGTA	ATCA	781		
663	Qy		GATTTCC	AGGC	TTTCT	CGATCG	GA	AAAG	ATTTGG	GGCAG	GAACAT	TAA	AAG	AGCCTTC	ATT	721	
782	Db		GATTTCC	AGGC	TTTCT	CGATCG	GA	AAAG	ATTTGG	GGCAG	GAACAT	TAA	AAG	AGCCTTC	ATT	841	
722	Qy		ATGACT	GAA	TTAC	ACTCAT	CTTT	TG	GAAG	AGTG	AGCA	AGCAG	TG	CGAGT	TTTTTTC	ACAG	781
842	Db		ATGACT	GAA	TTAC	ACTCAT	CTTT	TG	GAAG	AGTG	AGCA	AGCAG	TG	CGAGT	TTTTTTC	ACAG	901
782	Qy		CATCTTG	---	CTG	TGTTCA	ATTA	TTA	CCAT	CAC	AGCCTTTT	TAA	AAAA	ATCAT	CTTT	833	
902	Db		CATCTTG	CTG	GGG	GGTCA	ATTA	TTA	TTAC	CA	TCAC	CGG	CCCTTTT	TAA	CCAAA	ATCTTT	956

Search completed: June 30, 2005, 16:53:29
Job time : 6332.94 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 14:24:03 ; Search time 7762.11 Seconds
(without alignments)
10574.845 Million cell updates/sec

Title: US-09-945-173-1
Perfect score: 1694
Sequence: 1 agggaggagtggaagat.....aaacatgaattattgaacta 1694

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_cm.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1694	100.0	1694	6	AX465669
2	1680.6	99.2	3362	9	HSM807705
3	1680.6	99.2	3449	9	HSM807739
4	1666.8	98.4	1999	9	BC020832
5	1660.8	98.0	3352	9	HSM807810
6	1387	81.9	2550	9	AK025772
7	1022.2	60.3	65838	9	AC133474
8	1022.2	60.3	146010	9	AC009289
9	875.6	51.7	2064	10	BC050194
10	711	42.0	711	6	AX465671
11	488.2	28.8	3327	5	AJ720020
12	425	25.1	1064	5	BC073035
13	397	23.4	1040	5	BC078191
14	395.4	23.3	1291	5	AY391440
15	374.8	22.1	385	6	BD076273
16	330.8	19.5	155418	2	AC117754
17	321	18.9	321	6	AX906158
18	321	18.9	321	6	BD041691
19	301.8	17.8	245326	2	AC107288

20	301.8	17.8	266597	2	AC119722
21	223	13.2	223	6	CQ719549
22	139.2	8.2	66009	2	AC101250
23	135	8.0	119730	9	AC117472
24	135	8.0	188485	2	AC027454
25	102.2	6.0	155836	2	AC138153
26	102.2	6.0	174601	5	EX571724
27	102.2	6.0	180928	2	AC139727
28	92.6	5.5	66009	2	AC101250
29	81.6	4.8	87017	2	AC140043
30	77.6	4.6	29949	2	AC112488
31	77.6	4.6	179351	2	AC143587
32	67.2	4.0	188485	2	AC027454
33	66.8	3.9	919	3	CQ581943
34	66.8	3.9	934	3	AY070945
35	66.8	3.9	1087	6	AR508948
36	60	3.5	60	6	CQ540888
37	56.8	3.4	3048	6	CQ581942
38	56.8	3.4	3215	6	CQ591803
39	56.8	3.4	3888	6	CQ591806
40	56.8	3.4	8610	2	AC012956
41	56.8	3.4	177576	3	AC022349
42	56.8	3.4	179077	3	AC018489
43	56.8	3.4	300459	3	AE003503
44	51.8	3.1	1141	6	AX083744
45	51	3.0	146917	2	CR450801

ALIGNMENTS

RESULT 1
AX465669
LOCUS AX465669 1694 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO0218425.
ACCESSION AX465669
VERSION AX465669.1 GI:21899929
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Meyers, R.A.
TITLE 47324, a human g-protein and uses therefor
JOURNAL Patent: WO 0218425-A 1 07-MAR-2002;
Millennium Pharmaceuticals (US)
FEATURES
Location/Qualifiers
1..1694
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
19..729
/note="unnamed protein product"

CDS

1..1694
/codon_start=1
/protein_id="CAD42379.1"
/db_xref="GI:21899930"
/translation="MASLDRVKVLVLDGSGVGSILVHLICQNVLGNPSWTVGCSVD
VRVHDYKEGTPPEKTYIIEIWDVGGSGVSSSVKSTRVFNYSVNGIIFVHDLTNKKS
SQLRRWSLEALNRDLVPTGVLTVNGIDYDQEQFADNQIPLLVIGTKLDQDQTHETKRHEV
LRTAPLAEDFNPEINLDTCTNPRYLAAGSSNAVKLSRFPDKVIEKRYFLREGNQIPG
FDPKRFGAGTLKSLHYD"

ORIGIN

Query Match 100.0%; Score 1694; DB 6; Length 1694;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGAAGGAGTGGCAGAGTGGCGTCCCTGGATCGGTGAAGGTACTGTGTTGGGAGAC 60
DB 1 AGGGAAGGAGTGGCAGAGTGGCGTCCCTGGATCGGTGAAGGTACTGTGTTGGGAGAC 60
QY 61 TCAGGTGTTGGGAATCTTCCTAGTTCCTCTCTATGCCAAATCAAGTCTGCTGGGAAT 120

[illegible][illegible]

```

/tissue_type="human bone marrow"
/clone_lib="313 (synonym: hlcc2)". Vector pTriplEx2; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
3308. .3313
3332
polyA_signal
polyA_site
ORIGIN

```

Query Match	99.2%;	Score 1680.6;	DB 9;	Length 3362;
Best Local Similarity	99.8%;	Ident. No. 0;		
Matches 1683;	Conservative	0;	Mismatches 4;	Indels 0; Gaps 0;
Qy	1	AGGGAAGCAGCTGGCAAGATGCGTCCCTGGATCGGCTCAAGGTACTGCTGTGGGAGAC	60	
Db	10	AGGGAAGCAGTGGCAAGATGCGTCCCTGGATCGGCTCAAGGTACTGCTGTGGGAGAC	69	
Qy	61	TCAGTGTGTGGAAATCTTCTGTAGTCCATCTCTATGCGCAAAATCAAGTGTCTGGGAAT	120	
Db	70	TCAGTGTGTGGAAATCTTCTGTAGTCCATCTCTATGCGCAAAATCAAGTGTCTGGGAAT	129	
Qy	121	CCATCATGCACTGTGGGCTGCTCAGTGGATGTCAAGTTCATGATTACAAGAAAGGAACC	180	
Db	130	CCATCATGCACTGTGGGCTGCTCAGTGGATGTCAAGTTCATGATTACAAGAAAGGAACC	189	
Qy	181	CCAGAAGAGAGACCTACTACATAGAAATATGCGGATGTGGAGGCTCTGTGGGCAAGTGC	240	
Db	190	CCAGAAGAGAGACCTACTACATAGAAATATGCGGATGTGGAGGCTCTGTGGGCAAGTGC	249	
Qy	241	AGCAGCGTGAAAGCAACAAGAGCAGTATTTCTACAACTCCGTAAATGGTATTTATTTTCGTA	300	
Db	250	AGCAGCGTGAAAGCAACAAGAGCAGTATTTCTACAACTCCGTAAATGGTATTTATTTTCGTA	309	
Qy	301	CACGACTTAAACAATAAGAGTCTCCCAAAACTTGGCTGTGTGCTCATTTGAAGCTCTC	360	
Db	310	CACGACTTAAACAATAAGAGTCTCCCAAAACTTGGCTGTGTGCTCATTTGAAGCTCTC	369	
Qy	361	AACAGGGATTTGGTGCCAACTGGAGTCTTTGGTGACAAATGGGGATTTATGATCAAGAACAG	420	
Db	370	AACAGGGATTTGGTGCCAACTGGAGTCTTTGGTGACAAATGGGGATTTATGATCAAGAACAG	429	
Qy	421	TTTGCTGTATAACCAATAACCACTGTGTGGTAAATAGGGACTAAACTCGACAGATTTCATGAA	480	
Db	430	TTTGCTGTATAACCAATAACCACTGTGTGGTAAATAGGGACTAAACTCGACAGATTTCATGAA	489	
Qy	481	ACAAAGCCCATGAAGTTTTTAACTAGGACTGCTTCTCGCTGAGAGATTTCATCCAGNA	540	
Db	490	ACAAAGCCCATGAAGTTTTTAACTAGGACTGCTTCTCGCTGAGAGATTTCATCCAGNA	549	
Qy	541	GAATTAATTTGGACTGCACAAATCCACGGTACTTAGCTGCAAGTCTTCCAAATGCTGTC	600	
Db	550	GAATTAATTTGGACTGCACAAATCCACGGTACTTAGCTGCAAGTCTTCCAAATGCTGTC	609	
Qy	601	AAGCTCAGTAGGTTTTTGTATAAGGTCTATAGAGAGAGATCTTTTTTAAGAGAAGGTAAAT	660	
Db	610	AAGCTCAGTAGGTTTTTGTATAAGGTCTATAGAGAGAGATCTTTTTTAAGAGAAGGTAAAT	669	
Qy	661	CAGATTCAGGCTTCTCGATCGGAAAGATTTGGGCGAGGACAAATTAAGAGCCCTTCAT	720	
Db	670	CAGATTCAGGCTTCTCGATCGGAAAGATTTGGGCGAGGACAAATTAAGAGCCCTTCAT	729	
Qy	721	TATGACTGAATTTACTCATCTCTTTGGGAAGAGTGTAGCAAGCAGTGGCAGTTTTTTCACAGC	780	
Db	730	TATGACTGAATTTACTCATCTCTTTGGGAAGAGTGTAGCAAGCAGTGGCAGTTTTTTCACAGC	789	
Qy	781	TCATCTTGCTGTGTTCAATTTATACCATCACAGCCTTTTAAACAAATCATCTTAAAAATGC	840	
Db	790	TCATCTTGCTGTGTTCAATTTATACCATCACAGCCTTTTAAACAAATCATCTTAAAAATGC	849	
Qy	841	TACCTTTCAGGCTTACCTTTAAATGGAATAATGAAGGAGTGAACAATACGGGAGGTCCA	900	
Db	850	TACCTTTCAGGCTTACCTTTAAATGGAATAATGAAGGAGTGAACAATACGGGAGGTCCA	909	
Qy	901	AACTTTGTCCCTGTCTCTGTGTTCCCTTACCTTTCTGTCCCTGTGTATAGATTATGTAAA	960	

[illegible]

RESULT 3
HSM807739
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
CONSRPM
TITLE
JOURNAL

3449 bp mRNA linear PRI 30-AUG-2003
HSM807739
Homo sapiens mRNA; cDNA DKFP313M1333 (from clone DKFP313M1333).
BX647593
BX647593.1 GI:34366750
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3449)
Koehler K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
Fobo G., Han M. and Wiemann S.
The German Human cDNA Consortium
Direct Submission
Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764

	Neuherberg, GERMANY	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de; Sequenced by BMEP (Biomedical Research Center at the Heinrich-Heine-University, Duesseeldorf/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFP313M133) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.
FEATURES	Location/Qualifiers	
source	1..3449	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="DKFP313M133"	
	/tissue_types="human heart"	
	/clone_lib="313 (synonym: hlcc2). Vector pTriplEx2; host DH10B; sites SfIIA + SfIIB"	
	/dev_stage="adult"	
polya_signal	3311..3316	
polya_site	3336	
ORIGIN		
Query Match	99.2%; Score 1680.6;	DB 9; Length 3449;
Best Local Similarity	99.8%; Pred. No. 0;	
Matches 1683;	Conservative 0;	Mismatches 4; Indels 0; Gaps 0;
Qy	1 AGGGAAGCGCAGTGCACAGATGGCGTCCTCGATCGGTGAAGGTACTGGTGTTGGGAGAC	60
Db	13 AGGGAAGCGCAGTGCACAGATGGCGTCCTCGATCGGTGAAGGTACTGGTGTTGGGAGAC	72
Qy	61 TCAGGTGTGGGAATCTTCGTTAGTCCATCTCCTATGCCAAAATCAAGTGTGGGAAAT	120
Db	73 TCAGGTGTGGGAATCTTCGTTAGTCCAATCTCTATGCCAAATCAAGTGTGGGAAAT	132
Qy	121 CCATCATGGACTGTGGGTGCTCAGTGGATGTCAAGTTTCATGATTACAAGAAGGAACC	180
Db	133 CCATCATGGACTGTGGGTGCTCAGTGGATGTCAAGTTTCATGATTACAAGAAGGAACC	192
Qy	181 CCAGAAGAGAGACCTACTACATAGAATTATGGCATGTTCGAGGCTCTGTGGGCGAGTCC	240
Db	193 CCAGAAGAGAGACCTACTACATAGAATTATGGCATGTTCGAGGCTCTGTGGGCGAGTCC	252
Qy	241 AGCAGCGCTGAAAGCACAGAAGCAGTAGTATCTACAACTCCGTAATGTGTTATTTTCGTA	300
Db	253 AGCAGCGTGAAAGCACAGAAGCAGTAGTATCTACAACTCCGTAATGTGTTATTTTCGTA	312
Qy	301 CACGACTTACAAATAGAAGTCTCTCCAAAACHTTGGTGTGGTCAATTGGGAAGCTCTC	360
Db	313 CACGACTTACAAATAAGAAGTCTCTCCAAAACHTTGGTGTGGTCAATTGGGAAGCTCTC	372
Qy	361 AACAGGGATTTGGTGCCAACTGGAGTCTTGTGTGACAAATGGGATTTATGATCAAGAACAG	420
Db	373 AACAGGGATTTGGTGCCAACTGGAGTCTTGTGTGACAAATGGGATTTATGATCAAGAACAG	432
Qy	421 TTTTGCTGATAACCAAATACCACTGTGGTAAATAGGGACTAACTGGACCGAGATTCATGAA	480
Db	433 TTTTGCTGATAACCAAATACCACTGTGGTAAATAGGGACTAACTGGACCGAGATTCATGAA	492
Qy	481 ACAAGCGCCATGAAGTTTTTAACTAGGACTGCTTTCTTGCGCTAGGATTTCAATCCGAA	540
Db	493 ACAAGCGCCATGAAGTTTTTAACTAGGACTGCTTTCTTGCGCTAGGATTTCAATCCGAA	552
Qy	541 GAAATTAATTTGGAGCTGCACAAATCCACGGTACTTAGCTGCAGGTCTTCCAAATGCTGTC	600
Db	553 GAAATTAATTTGGAGCTGCACAAATCCACGGTACTTAGCTGCAGGTCTTCCAAATGCTGTC	612
Qy	601 AAGCTCAGTAGGTTTTTTTGATAAAGGTCATAGAGAAGAGATATCTTTTTAAGAAGAGTAAAT	660
Db	613 AAGCTCAGTAGGTTTTTTTGATAAAGGTCATAGAGAAGAGATATCTTTTTAAGAAGAGTAAAT	672


```
Db 781 TGTGTTCAATTATACCATCACAGCCTTTTAAACAAATCATCTTAAATGCTACCCCTTCA 840
QY 850 GCCTTACCTTTAATGGAATAATGAAGAGTGCACAAATACGGGAGGTCCAAACTTTGTC 909
Db 841 GCCTTACCTTTAATGGAATAATGAAGAGTGCACAAATACGGGAGGTCCAAACTTTGTC 900
QY 910 CCGTTCCTGCTGCTCCTTACCTTTCTGTCCTGCTGTATAGATTATGTAAAGCCTTGTG 969
Db 901 CCGTTCCTGCTGCTCCTTACCTTTCTGTCCTGCTGTATAGATTATGTAAAGCCTTGTG 960
QY 970 TAAATATGAGATGTTGTCAAAATGATGACGTAATGAGCAATGACAGTGTACTGCAGAGA 1029
Db 961 TAAATATGAGATGTTGTCAAAATGATGACGTAATGAGCAATGACAGTGTACTGCAGAGA 1020
QY 1030 AAATTTACTCTTGCTAGAACTGGAGGTTTTTATGGGCTGTGAATTTTCCCACTCAT 1089
Db 1021 AAATTTACTCTTGCTAGAACTGGAGGTTTTTATGGGCTGTGAATTTTCCCACTCAT 1080
QY 1090 TGTGTAAGAGCTTAAATTAAGTACTTCAAAACGATCTCCATTCGTTTACCTCTTGAGGG 1149
Db 1081 TGTGTAAGAGCTTAAATTAAGTACTTCAAAACGATCTCCATTCGTTTACCTCTTGAGGG 1140
QY 1150 GAAAGGCTTTGTTAACAGCCCTGAGTGTGTACCCCAACAACTCTGTCAATTTCAA 1209
Db 1141 GAAAGGCTTTGTTAACAGCACTGAGTGTGTACCCCAACAACTCTGTCAATTTCAA 1200
QY 1210 GATGCAAAATGGTGTATTAATTTGCTCCACATTGTCAACACAGGAATGCTTAATA 1269
Db 1201 GATGCAAAATGGTGTATTAATTTGCTCCACATTGTCAACACAGGAATGCTTAATA 1260
QY 1270 TAGCAACCTTGTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1329
Db 1261 TAGCAACCTTGTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
QY 1330 ACTAATAGCCAGAGTTAAATATATAAATACAAATTAATAATACATAGAGAACAGCAATACCA 1389
Db 1321 ACTAATAGCCAGAGTTAAATATATAAATACAAATTAATAATACATAGAGAACAGCAATACCA 1380
QY 1390 GAAAAAAGAAATTCGTGTAATAATGATGTGAAAAATGACAGTCCCTCACTCTTAAGGTT 1449
Db 1381 GAAAAAAGAAATTCGTGTAATAATGATGTGAAAAATGACAGTCCCTCACTCTTAAGGTT 1440
QY 1450 GCTGCTATATACAGTCTAGGTTTTCTGTTTGGAAATAGGTAGGTAATAATCTAAGACCTG 1509
Db 1441 GCTGCTATATACAGTCTAGGTTTTCTGTTTGGAAATAGGTAGGTAATAATCTAAGACCTG 1500
QY 1510 CACAAGGCGAGTGAGAGACATTTACAGCCCTCTCTCTATTTGTTTTTTTAAAGGAAAAGTC 1569
Db 1501 CACAAGGCGAGTGAGAGACATTTACAGCCCTCTCTCTATTTGTTTTTTTAAAGGAAAAGTC 1560
QY 1570 AACTCTGAAATGTCCTTAGCTATTAATCGAAAAATTAAGAAATTAATTTCTGTGTCAACA 1629
Db 1561 AACTCTGAAATGTCCTTAGCTATTAATCGAAAAATTAAGAAATTAATTTCTGTGTCAACA 1620
QY 1630 ATGTTATTTAGGAGAGAGTAAATAAGTTTCCACAGACACACAAAAATGAATTAAT 1687
Db 1621 ATGTTATTTAGGAGAGAGTAAATAAGTTTCCACAGACACACATTTACATGAATTAAT 1678
```

RESULT 5

```
HSM807810
LOCUS 3352 bp mRNA linear PRI 30-AUG-2003
DEFINITION Homo sapiens mRNA; cDNA DKFp3130171 (from clone DKFp3130171).
ACCESSION BX647664
VERSION BX647664.1 GI:34366821
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3352)
AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
```

```
Wiemann, S.
The German Human cDNA Consortium
Direct Submission
Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFp3130171) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
    source          1..3352
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="DKFp3130171"
                    /tissue_type="human fetal skin"
                    /clone_lib="313 (synonym: hlcc2). Vector pTriplex2; host
                    DH10B; sites SfiIA + SfiIB"
                    /dev_stage="fetal"
                    polyA_signal 3296..3301
                    polyA_site 3321
ORIGIN
    Query Match      98.0%; Score 1660.8; DB 9; Length 3352;
    Best Local Similarity 99.5%; Pred. No. 0;
    Matches 1676; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 5 AAGGCAGTCGCAAGATGCGCTCCCTGGATCGGGTGAAGGTACTGTGTGGGAGACTCAG 64
Db 1 AAGGCAGTCGCAAGATGCGCTCCCTGGATCGGGTGAAGGTACTGTGTGGGAGACTCAG 60
QY 65 GTGTTGGGAAATCTTCGTTAGTCCATCTCTATGCCAAAATCAAGTCTGGGAAATCCAT 124
Db 61 GTGTTGGGAAATCTTCGTTAGTCCATCTCTATGCCAAAATCAAGTCTGGGAAATCCAT 120
QY 125 CATGACTCTGGCTGCTCAGTGGATGTCAGAGTTCATGATTAACAAGAGAACCCAG 184
Db 121 CATGACTCTGGCTGCTCAGTGGATGTCAGAGTTCATGATTAACAAGAGAACCCAG 180
QY 185 AAGAGAAGACCTACTACATAGAAATTAATGGATGTTGGAGGCTCTGTGGGCAAGTCCAGCA 244
Db 181 AAGAGAAGACCTACTACATAGAAATTAATGGATGTTGGAGGCTCTGTGGGCAAGTCCAGCA 240
QY 245 GCGTGAAGAGCACAGAGCAGTATTTCTACAACTCCGTAATGATTAATTTTGGTACAG 304
Db 241 GCGTGAAGAGCACAGAGCAGTATTTCTACAACTCCGTAATGATTAATTTTGGTACAG 300
QY 305 ACTTAACAAATGAAGAGTCTCTCCAAAATCTTGGCTGTTGGTCAITTTGGAAGCTCTCAACA 364
Db 301 ACTTAACAAATGAAGAGTCTCTCCAAAATCTTGGCTGTTGGTCAITTTGGAAGCTCTCAACA 360
QY 365 GGGATTTGGTGCACACTGGAGTCTTGGTGACAAATGGGGATTATGATCAAGAACAGTTTG 424
Db 361 GGGATTTGGTGCACACTGGAGTCTTGGTGACAAATGGGGATTATGATCAAGAACAGTTTG 420
QY 425 CTGATTAACAAATACCACTGTTGTTAATAGGGACTAACTGGACCCAGATTCATGAACAA 484
Db 421 CTGATTAACAAATACCACTGTTGTTAATAGGGACTAACTGGACCCAGATTCATGAACAA 480
QY 485 AGCGCATCAAGTTTTAACTAGGACTGCTTTTCTCGCTGAGGATTTCAATCCAGAGAAA 544
Db 481 AGCGCATCAAGTTTTAACTAGGACTGCTTTTCTCGCTGAGGATTTCAATCCAGAGAAA 540
QY 545 TTAATTTGAGCTGCACAAATCCACGGTACTTAGCTCGAGGTTCTTCCAAATGCTGTCAAGC 604
Db 541 TTAATTTGAGCTGCACAAATCCACGGTACTTAGCTCGAGGTTCTTCCAAATGCTGTCAAGC 600
```

Wiemann, S.

The German Human cDNA Consortium

Direct Submission

Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764

Neuherberg, GERMANY

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

This clone (DKFp3130171) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further

information about the clone and the sequencing project is available

at http://mips.gsf.de/proj/cDNA/.

Location/Qualifiers

1..3352

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFp3130171"

/tissue_type="human fetal skin"

/clone_lib="313 (synonym: hlcc2). Vector pTriplex2; host

DH10B; sites SfiIA + SfiIB"

/dev_stage="fetal"

polyA_signal 3296..3301

polyA_site 3321

ORIGIN

Query Match

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1676; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```
Qy 605 TCAGTAGGTTTTTGATAGAGTCTATAGAGAGAGATATCTTTTTTAAGAGAGAGTAAATCAGA 664
Db 601 TCAGTAGGTTTTTGATAGAGTCTATAGAGAGAGATATCTTTTTTAAGAGAGAGTAAATCAGA 660
Qy 665 TTCCAGGCTTTCTGATCGGAAGATTTGGGGCAGGAAACATTAAGAGCCTTCATTATG 724
Db 661 TTCCAGGCTTTCTGATCGGAAGATTTGGGGCAGGAAACATTAAGAGCCTTCATTATG 720
Qy 725 ACTGAATTAACACTCATCTTTTGAAGAGTGAAGCAGTGGCGAGTGTTCACAGCTCAT 784
Db 721 ACTGAATTAACACTCATCTTTTGAAGAGTGAAGCAGTGGCGAGTGTTCACAGCTCAT 780
Qy 785 CTTCGCTGTTCATTATTACCATCACAGCCTTTTAAACAAATCATCTTAAATGCTACC 844
Db 781 CTTCGCTGTTCATTATTACCATCACAGCCTTTTAAACAAATCATCTTAAATGCTACC 840
Qy 845 CTTGAGCCTTACCTCTTAATGGAAGATGAAGGAAGTGAACATACGGGAGGTCCAAACT 904
Db 841 CTTGAGCCTTACCTCTTAATGGAAGATGAAGGAAGTGAACATACGGGAGGTCCAAACT 900
Qy 905 TTGTCCTGTCTCTGTCTTACCTTTCTGTCCTCTGTGTATAGATTATGTAAGGCC 964
Db 901 TTGTCCTGTCTCTGTCTTACCTTTCTGTCCTCTGTGTATAGATTATGTAAGGCC 960
Qy 965 TTGTGTAATATGAGATGTTGTCAAAATGATGCAAGTAAATGAGCAATGACAGTGTACTGC 1024
Db 961 TTGTGTAATATGAGATGTTGTCAAAATGATGCAAGTAAATGAGCAATGACAGTGTACTGC 1020
Qy 1025 AGAGAAATTTACTCTTCCTAGAACCTGGAGGTTTTTATGGGCTGTGTAATTTCCACCA 1084
Db 1021 AGAGAAATTTACTCTTCCTAGAACCTGGAGGTTTTTATGGGCTGTGTAATTTCCACCA 1080
Qy 1085 CTCATTCTGAAAGCTTAATTAAGTACTTCAAAACGTAATCTCCATTGTTTACCTCTT 1144
Db 1081 CTCATTCTGAAAGCTTAATTAAGTACTTCAAAACGTAATCTCCATTGTTTACCTCTT 1140
Qy 1145 GAGGGAAACGGTCTGTTTAAACAGCCCTGAGTGTCTACCCCAACAACTCTCTGCTATT 1204
Db 1141 GAGGGAAACGGTCTGTTTAAACAGCCCTGAGTGTCTACCCCAACAACTCTCTGCTATT 1200
Qy 1205 TCAAGATGCAAAATGGTGTATTTAATGCTCCACCATTTGTCACACACA-GGAATGCC 1263
Db 1201 TCAAGATGCAAAATGGTGTATTTAATGCTCCACCATTTGTCACACACAGGGAATGCC 1260
Qy 1264 TAATAATAGCAACCTTGTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1323
Db 1261 TAATAATAGCAACCTTGTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
Qy 1324 AGCGGACTAATAGCCAGAGTTAAATATAATAACAAATTAATAATACATAGAGAACAGCA 1383
Db 1321 AGCGGACTAATAGCCAGAGTTAAATATAATAACAAATTAATAATACATAGAGAACAGCA 1380
Qy 1384 ATACCAGAAAAGAAATCTGTTAAATGATGTGAAAATTTGACAGCTCCCTCAGCTTT 1443
Db 1381 ATACCAGAAAAGAAATCTGTTAAATGATGTGAAAATTTGACAGCTCCCTCAGCTTT 1440
Qy 1444 AAGTTCTCTGTATATACAGTCTAGGTTTCTGTTTGAATAGTGGGTAAATCTAA 1503
Db 1441 AAGTTCTCTGTATATACAGTCTAGGTTTCTGTTTGAATAGTGGGTAAATCTAA 1500
Qy 1504 GACCTGCAAGGGCAGTGAGAGACATTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1563
Db 1501 GACCTGCAAGGGCAGTGAGAGACATTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560
Qy 1564 AAGTCAACTCTGAAATGTCCTTAGTATAATCAGAAATAGTAATATTATCTGTG 1623
Db 1561 AAGTCAACTCTGAAATGTCCTTAGTATAATCAGAAATAGTAATATTATCTGTG 1620
Qy 1624 TCAACATGTTATTATGAGAGAGTAAATAAGTTCCACAGCAACACAAAAACATGAA 1683
Db 1621 TCAACATGTTATTATGAGAGAGTAAATAAGTTCCACAGCAACACATTTACATGAA 1680
Qy 1684 TTAT 1687
```

```
Db 1681 TTAT 1684
```

RESULT 6

AK025772
LOCUS Homo sapiens cDNA: FLJ22119 fis, clone HEP18852.
DEFINITION Homo sapiens cDNA: FLJ22119 fis, clone HEP18852.
ACCESSION AK025772
VERSION AK025772.1 GI:10438390
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Kawabata, A., Hikiji, T.; Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

FEATURES

source

Location/Qualifiers

1. .2550

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="HEP18852"

/cell_line="HepG2"

/cell_type="hepatoma"

/clone_lib="HEP"

/note="Cloning vector pME18SFL3"

ORIGIN

Query Match

Best Local Similarity

Matches 1390; Conservative

81.9%; Score 1387; DB 9; Length 2550;

99.6%; Pred. No. 0;

Mismatches 5; Indels 0; Gaps 0;

Qy 293 TTTTCGTACAGCTTAACAAATAGAGAGTCTCCCAAACTTCGGTGGTCATTGG 352

Db 1 TTTTCGTACAGCTTAACAAATAGAGAGTCTCCCAAACTTCGGTGGTCATTGG 60

Qy 353 AAGCTCTCAACAGAGTTTGGTGCCAACTGGAGTCTTTGGTGACAAATGGGATTTATGATC 412

Db 61 AAGCTCTCAACAGAGTTTGGTGCCAACTGGAGTCTTTGGTGACAAATGGGATTTATGATC 120

Qy 413 AAGAACATGTTTGTGTATACCAATACCACTGTTTGGTAATAGGAGCTAAATCTGGACCA 472

Db 121 AAGAACATGTTTGTGTATACCAATACCACTGTTTGGTAATAGGAGCTAAATCTGGACCA 180

Qy 473 TTCTATGAACAAAGCGGCATGAAGTTTAACTAGGACTGCTTTCTGCTGAGGATTTCA 532

Db 181 TTCTATGAACAAAGCGGCATGAAGTTTAACTAGGACTGCTTTCTGCTGAGGATTTCA 240

Qy 533 ATCCAGAGAAATTAATTTGGAGCTGCACAAATCCACGCTACTTACGCTGCAGGTTCTTCCA 592

Db 241 ATCCAGAGAAATTAATTTGGAGCTGCACAAATCCACGCTACTTACGCTGCAGGTTCTTCCA 300

QY 593 ATGCTGTCAAGCTCAGTAGGTTTTTTTGAATAGGTCTATAGAGAGAGATCTTTTTTAAGAG 652
DB |||||
301 ATGCTGTCAAGCTCAGTAGGTTTTTTGATAGGTCTATAGAGAGAGATCTTTTTTAAGAG 360
QY 653 AAGGTAATCAGATTCAGGGCTTCTCGATCGGAAAAGATTGTTGGGCGAGAACATTTAAGA 712
DB |||||
361 AAGGTAATCAGATTCAGGGCTTCTCGATCGGAAAAGATTGTTGGGCGAGAACATTTAAGA 420
QY 713 GCCTTCATTATGACTGAATTTACACTCATCTCTTTGGAGAGCTGAGCAGAGTGGCAGTTT 772
DB |||||
421 GCCTTCATTATGACTGAATTTACACTCATCTCTTTGGAGAGCTGAGCAGAGTGGCAGTTT 480
QY 773 TTCACAGCTCATCTTGCTGTGTTCAATTTATACCATCACAGCCTTTTAAACAAATCATCT 832
DB |||||
481 TTCACAGCTCATCTTGCTGTGTTCAATTTATACCATCACAGCCTTTTAAACAAATCATCT 540
QY 833 TAAATGCTACCTTCAGCCTTACCCCTTTAATGGAATAATGAAAGGAAGTGACAATACGG 892
DB |||||
541 TAAATGCTACCTTCAGCCTTACCCCTTTAATGGAATAATGAAAGGAAGTGACAATACGG 600
QY 893 GAGGTCCTAACTTTGTCCTGTTCTCTGTTCTCTTACCTTTCTGTCCTCTGTATAGAT 952
DB |||||
601 GAGGTCCTAACTTTGTCCTGTTCTCTGTTCTCTTACCTTTCTGTCCTCTGTATAGAT 660
QY 953 TATGTAAGACCTTGTGTAATATGAGATGTTGTCAAAATGATGTCAGTAATGAGCAATG 1012
DB |||||
661 TATGTAAGACCTTGTGTAATATGAGATGTTGTCAAAATGATGTCAGTAATGAGCAATG 720
QY 1013 ACAGTGTACTGCAGAGAAAATTTACTCTTCCTCTAGAACCTGAGGGTTTTTATGGGTCTGT 1072
DB |||||
721 ACAGTGTACTGCAGAGAAAATTTACTCTTCCTCTAGAACCTGAGGGTTTTTATGGGTCTGT 780
QY 1073 AATTTTCCACACTCATCTTGCTGAAGCTTAAATTAAGTACTTCAAAACGTTATCTCCATG 1132
DB |||||
781 AATTTTCCACACTCATCTTGCTGAAGCTTAAATTAAGTACTTCAAAACGTTATCTCCATG 840
QY 1133 TTTTACCTCTTCGAGGGGAACGGTCTTGTTAAACAGCCCTGAGTGTGTCTACCCCAACAA 1192
DB |||||
841 TTTTACCTCTTCGAGGGGAACGGTCTTGTTAAACAGCCCTGAGTGTGTCTACCCCAACAA 900
QY 1193 TCTCTGTCAATTTTCAAGATGCAAAAATGGTGTATTATTAATTTGTCTCACCAATTTGCACAC 1252
DB |||||
901 TCTCTGTCAATTTTCAAGATGCAAAAATGGTGTATTATTAATTTGTCTCACCAATTTGCACAC 960
QY 1253 ACAGGAATGCTTAATATACACACCTTGTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1312
DB |||||
961 ACAGGAATGCTTAATATACACACCTTGTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
QY 1313 GTGACTGGAAGAGCGGACTTAATAGCCAGAGTTAAATATAAATACAAATTAATAATACAT 1372
DB |||||
1021 GTGACTGGAAGAGCGGACTTAATAGCCAGAGTTAAATATAAATACAAATTAATAATACAT 1080
QY 1373 AGAGAACAGCAATACCAAGAAAAAAGAAATTTCTGGTAAATATGATGTGAAAAATTTGACAGCT 1432
DB |||||
1081 AGAGAACAGCAATACCAAGAAAAAAGAAATTTCTGGTAAATATGATGTGAAAAATTTGACAGCT 1140
QY 1433 CCTCACTCTTAGGTTGCTGCTATATACAGCTTAGGTTTCTGTTTGGAAATAGGTAGG 1492
DB |||||
1141 CCTCACTCTTAGGTTGCTGCTATATACAGCTTAGGTTTCTGTTTGGAAATAGGTAGG 1200
QY 1493 GTAAATCTAAGACCTTCACAAAGGCGAGTGAGAGACATTTTACAGCCCTCTCTCTATTGTT 1552
DB |||||
1201 GTAAATCTAAGACCTTCACAAAGGCGAGTGAGAGACATTTTACAGCCCTCTCTCTATTGTT 1260
QY 1553 TTTTTTAAGAAAAAGTCAACTCTCGAATGTCCTTTAGCTTATATACGAAAAATTAAGAAT 1612
DB |||||
1261 TTTTTTAAGAAAAAGTCAACTCTCGAATGTCCTTTAGCTTATATACGAAAAATTAAGAAT 1320
QY 1613 ATTATTTCTGTCAACAAATGTTATTTATGGAGAGAGTAAATAAGTTTCCACAGCCACAC 1672
DB |||||
1321 ATTATTTCTGTCAACAAATGTTATTTATGGAGAGAGTAAATAAGTTTCCACAGCCACAC 1380
QY 1673 AAAAAACATGAATTAT 1687

Db 1381 ATTACATGAATTAT 1395

RESULT 7
AC133474LOCUS
DEFINITIONAC133474 65838 bp DNA linear PRI 03-DEC-2002
Homo sapiens 3 BAC RP11-719N16 (Roswell Park Cancer Institute Human
BAC library) complete sequence.AC133474
AC133474.2 GI:23334664VERSION
KEYWORDSSOURCE
HTG.ORGANISM
Homo sapiens (human)Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.REFERENCE
AUTHORS1 (bases 1 to 65838)
Muzny,D.M., Adams,C., Amaratunga,H.C., Are,J.R., Ayele,M., Banks,T.,
Alsbrooks,S.L., Bimaga,K., Blankenburg,K., Bonnin,D.,
Barbaria,J., Benton,J., Brieve,M., Brown,E., Brown,M., Bryant,N.P.,
Bouck,J., Bowls,S., Burrell,C., Burrell,K.L., Byrd,N.C.,
Buhay,C., Burch,P., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Carron,T.F., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Chen,G., Chen,R., Cox,C., Coyie,M.D., Dathorne,S.R., David,R.,
Cleveland,C.D., Cox,C., Coyie,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Deltney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabris,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Liu,W.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenwo,S., Ogih,M., Okunolu,G., Oragunye,N.,
Oviedo,A., Pace,A., Payton,B., Peery,J., Perez,L., Peters,M.,
Picks,R., Primus,E., Pu,L.L., Quiles,M., Ruiz,S., Saverly,G., Scherer,S.,
Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S.,
Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E.,
Sonaika,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasequez,L.,
Vera,V., Vallalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 65838)

Worley,K.C.

Direct Submission

Submitted (13-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 65838)

Worley,K.C.

Direct Submission

Submitted (27-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 65838)


```
repeat_region complement(561..1188)
repeat_region /rpt_family="L1MC5"
repeat_region /rpt_family="AT_rich"
repeat_region complement(3296..3585)
repeat_region /rpt_family="AluSg"
repeat_region complement(4016..4190)
repeat_region /rpt_family="MIR"
repeat_region complement(4310..4499)
repeat_region /rpt_family="MERSA"
repeat_region 4835..4876
repeat_region /rpt_family="(TG)n"
repeat_region complement(5956..6090)
repeat_region /rpt_family="MIR"
repeat_region 7328..7381
repeat_region /rpt_family="AT_rich"
repeat_region 8624..8678
repeat_region /rpt_family="AT_rich"
repeat_region 8972..9270
repeat_region /rpt_family="AluY"
repeat_region complement(9782..10144)
repeat_region /rpt_family="MER47A"
repeat_region 12654..12673
repeat_region /rpt_family="(TG)n"
repeat_region 13315..13378
repeat_region /rpt_family="AT_rich"
repeat_region 13379..13659
repeat_region /rpt_family="AluSx"
repeat_region 13660..13697
repeat_region /rpt_family="(TAAA)n"
repeat_region 14020..14170
repeat_region /rpt_family="FRAM"
repeat_region 14390..14488
repeat_region /rpt_family="LTR29B"
repeat_region 14464..14842
repeat_region /rpt_family="MER39"
repeat_region complement(15504..15690)
repeat_region /rpt_family="MERSA"
repeat_region complement(16098..16251)
repeat_region /rpt_family="HAL1"
repeat_region 16303..16430
repeat_region /rpt_family="L2"
repeat_region 16571..16727
repeat_region /rpt_family="FRAM"
repeat_region 16733..17021
repeat_region /rpt_family="AluJo"
repeat_region 17022..17048
repeat_region /rpt_family="AT_rich"
repeat_region complement(17834..17999)
repeat_region /rpt_family="MLT1H"
repeat_region complement(18178..18280)
repeat_region /rpt_family="MLT1H"
repeat_region complement(19886..20064)
repeat_region /rpt_family="MIR"
repeat_region complement(20532..20662)
repeat_region /rpt_family="L2"
repeat_region complement(20710..20822)
repeat_region /rpt_family="L2"
repeat_region 21289..21309
repeat_region /rpt_family="AT_rich"
repeat_region 21318..21381
repeat_region /rpt_family="MIR"
repeat_region complement(21377..21489)
repeat_region /rpt_family="MERSA"
repeat_region complement(21515..22019)
repeat_region /rpt_family="MLT1D"
repeat_region complement(23539..24609)
repeat_region /rpt_family="MER11C"
repeat_region complement(24828..24935)
repeat_region /rpt_family="L2"
repeat_region 27466..27711
repeat_region /rpt_family="MIR"
repeat_region 28480..28514
```

```
repeat_region /rpt_family="(CCTA)n"
repeat_region complement(28668..28796)
repeat_region /rpt_family="MIR"
repeat_region complement(29313..29982)
repeat_region /rpt_family="L1PA9"
repeat_region 29986..30244
repeat_region /rpt_family="L1ME"
repeat_region 31126..31164
repeat_region /rpt_family="AT_rich"
repeat_region complement(31202..31562)
repeat_region /rpt_family="THE1C"
repeat_region complement(31609..31976)
repeat_region /rpt_family="MER57B"
repeat_region 31980..32021
repeat_region /rpt_family="(TG)n"
repeat_region 32637..32797
repeat_region /rpt_family="MIR"
repeat_region 32905..32948
repeat_region /rpt_family="CT-rich"
repeat_region complement(33347..33529)
repeat_region /rpt_family="MER91A"
repeat_region 33642..33695
repeat_region /rpt_family="AT_rich"
repeat_region complement(34118..34426)
repeat_region /rpt_family="AluSg"
repeat_region 34662..34971
repeat_region /rpt_family="AluSp"
repeat_region 34972..34996
repeat_region /rpt_family="(TAA)n"
repeat_region 36739..37053
repeat_region /rpt_family="AluJb"
repeat_region 37054..37093
repeat_region /rpt_family="(GAAA)n"
repeat_region 37581..37607
repeat_region /rpt_family="AT_rich"
repeat_region 37647..37692
repeat_region /rpt_family="AT_rich"
repeat_region 37905..38174
repeat_region /rpt_family="THE1B"
repeat_region 38229..38317
repeat_region /rpt_family="CT-rich"
repeat_region complement(38318..38683)
repeat_region /rpt_family="MLT1B"
repeat_region 38684..38708
repeat_region /rpt_family="CT-rich"
repeat_region complement(38888..39013)
repeat_region /rpt_family="MERSA"
repeat_region 39074..39186
```

Query Match 60.3%; Score 1022.2; DB 9; Length 146010;
Best Local Similarity 99.7%; Pred. No. 2e-240;
Matches 1024; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy 661 CAGATTCAGGCTTTCCTGATCGGAAAAGATTTGGGCGAGGAACATTAAAGAGGCTTCAT 720
Db 11578 CAGATTCAGGCTTTCCTGATCGGAAAAGATTTGGGCGAGGAACATTAAAGAGGCTTCAT 11637
Qy 721 TATGACTGAATTACACTCATCTCTTGAAGAGTCAGCAAGCAGTCGCGAGTCTTTTTCACAGC 780
Db 11638 TATGACTGAATTACACTCATCTCTTGAAGAGTCAGCAAGCAGTCGCGAGTCTTTTTCACAGC 11697
Qy 781 TCATCTTGTGTTGTTCAATTATTATTCACATCAGGCTTTTAAACAAATCATCTTAAATGC 840
Db 11698 TCATCTTGTGTTGTTCAATTATTATTCACATCAGGCTTTTAAACAAATCATCTTAAATGC 11757
Qy 841 TACCCTTCAGGCTTACGCTTTTAAATGGAATAATGAAGAGTGAATACACGGAGGTCCA 900
Db 11758 TACCCTTCAGGCTTACGCTTTTAAATGGAATAATGAAGAGTGAATACACGGAGGTCCA 11817
Qy 901 AACTTTGTCCCTGTTCTCTGTGTTCTTACCTTTCTGTCCTGTGTATAGATTATGATAA 960
Db 11818 AACTTTGTCCCTGTTCTCTGTGTTCTTACCTTTCTGTCCTGTGTATAGATTATGATAA 11877
```



```
misc_feature      FSDKRFGGALKNFHCD"
21..464
/ gene="Rab13"
/ note="RAB; Region: Rab subfamily of small GTPases"
/ db_xref="CDD:smart00175"

ORIGIN
Query Match      51.7%; Score 875.6; DB 10; Length 2064;
Best Local Similarity 79.1%; Pred. No. 3.2e-204;
Matches 1201; Conservative 0; Mismatches 279; Indels 38; Gaps 12;

Qy 17 AGATGGCGTCCCTCGATCGGCTGAAGTACTGCTGGGAGACTCAGGCTGTCGGAAAT 76
Db 1 AGATGGCGTCCCTCGATCGGCTGAAGGTTCTGGGTGGGAGACTCAGGCTGTCGGAAAT 60

Qy 77 CTTCGTTAGTCCATCTCTCTATGCCAAATCAAGTGTCTGGGAAATCCATCATGTGACTGTGG 136
Db 61 CTTCAGTCCGTCATCTCTCTGTCGCCACAATCAAGTGTCTAGGAAATCCGTCATGTGACTGTGG 120

Qy 137 GCTGCTCAGTGGATGTGAGTTCATGATTAACAAGAGGAACCCGAGAGAGAACCT 196
Db 121 GCTGCTCGGTGGATATCAGAGTTCATGACTACAAGAGGAGGACCCCTGAAGAGAGACAT 180

Qy 197 ACTACATAGATTAATGGATGTTCGAGGCTCTGGGCGAGTCCGAGCAGCGTGAAGACA 256
Db 181 ACTATATAGAACTATGGATGTTCGAGGCTCTGGGCGAGTCCGAGCAGTGTGAAGACA 240

Qy 257 CAAGAGCAGTATTTCTACAACTCCGTAAATGGTATTTATTTTCGTACACGACTTAAACAATA 316
Db 241 CAAGAGCGGTGTTCTACACTCTGTAAACGACATCTTTAGTACATGACTTAAACAATA 300

Qy 317 AGAAGTCTCTCCAAAACCTTGGCTGCTGGTCAATTTGGAAAGCTCTCAACAGGAGATTGGTGC 376
Db 301 AGAAGTCACTCAAAAACCTTATATCGCTGGTCAATTTGGAAAGTCTCAATCGGAGTGCAGTTC 360

Qy 377 CNACTGAGTCTGGTGACAAATGGGATTTATGATCAAGACAGTCTTGGCTGATTAACCAA 436
Db 361 CAACCGGAGTCTGGTGACAAATGGGATTTATGATCAAGACAGTCTTGGCTGATTAACCAA 420

Qy 437 TACCACATGTTGGTAAATAGGAGTAACTGGACACAGATTCAATGAACAAAGCGCATGAAG 496
Db 421 TCCACATGTTGGTAAATAGGAGTAACTGGACACAGATTCAATGAACAAAGCGCATGAAG 480

Qy 497 TTTTAACTAGGACTGCTTTCTCGCTGAGGATTTCAATCCAGAAAGAAATTAATTTGAGCT 556
Db 481 TTTTAACTAGGACTGCTTTCTCGCTGAGGACTTCAATGCTGAAGAGATTAAATTTGAGT 540

Qy 557 GCACAAATCCACGCTACTAGTACAGTCTTCCATATGCTGTCAAGCTCAGTAGGTTTT 616
Db 541 GCACAAATCCACGCTCTCAGCTCAGGCTCTCCTCAATGCTGTCAAGCTCAGTAGATTTT 600

Qy 617 TTGATAAGGTCATAGAGAGAGATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTC 676
Db 601 TTGATAAGGTTATAGAGAGAGATTTTAAAGAGAGGTAATCAGATTCCAGGCTTTC 660

Qy 677 CTGATCGGAAAGATTTGGGCGAGGAACTTAAAGAGCTTTCAATATGACTGAATTAAC 736
Db 661 CTGATCGGAGAGGTTTGGAGGAGGCGCACTGAAGAACTTCCACTGTGACTGAGTGCC 720

Qy 737 TCATCTTTGGAAGAGTGAGCAAGCAGTGGCAGTGTTCACAGCTCATCTTGTGCTGTTC 796
Db 721 TCAGCC-CTGGAAGAGCAAGGGAACAGCAGCAGTGTCCACAGCTTCTCTTGTGCTGTAG 779

Qy 797 AATTAATACCATCAGCCTTTTAAACAATCATCTTAAATGCTACCCITCAGCCTTAC 856
Db 780 GGACATTAACCTTCTCAGGC-TTGGGAAAAACATCTCAAAATGCTA-----CCTCAC 830

Qy 857 CTTTAAATGGAATAATGAAGAGAGTGAACAATACGGAGGTCCTCAAACTTTGCTCCTGTTC 916
Db 831 CTTCTCTTGGAAATTTGAAGAGAGCAGCGCATGGAGGTCCTAGCTCTGTGCTGCCATCTC 890

Qy 917 TCTGTGTTCCCTTACCTTCTGTCTCCTGTGTATAGATTATGTAAGAGCCTGTGTGTAATAT 976
Db 1 TCTGTGTTCCCTTACCTTCTGTGTTGGAGTACTGCTGTGTTGGAGACTCAGGTTGTTGGAAATCT 60

Db 891 TGTA---TCCTCACCCCTGCTCCTGATTATAGCTTATGTAAAGCCTTGTGAGCAT 947
Qy 977 GAGATGTTGTCAAAATGATGAGTAAATGAGCAATGACAGTGTACTGTCAGAGAAAATTTA 1036
Db 948 AAAGGCTTGTCAAAACCCATCAGTAAAGAACTGAGAGACA-AGGCTGTCAGCGAAGGCTGA 1006

Qy 1037 CTCTTGCTAGAACTGAGGCTTTTATGCGTCTGTAAATTTTCCACACTCATTTGCTGAA 1096
Db 1007 CTCTTACTTGGAAACAGCGGG-TTTTACAGCGCTGCGATTTTCCACACTTTGCTGCTGAA 1065

Qy 1097 AGCTTAATTAAGTACTTCAAAAACGATCTCTCAATGTTGTTTACCTTCTTGAGGGCAACGGT 1156
Db 1066 GGCTTAATTAAGTACTTCAAAAACATATCTGTGTTGTTTA-CTTCTGAGGGGAAGGT 1124

Qy 1157 CTGTGTAACGAGCTCAGTGTGTACCCCAAAACAATCTCTGTCTATTTTCAAGATGCAA 1216
Db 1125 CATGTTAATCAGCGCTGAGTTATCCACCCCAACCATCCCTGCCATTTTAAAGAGAGCA 1184

Qy 1217 AATGGTGTATTTAATTTGTTCTCCACATTTGTCAACACAGGAATGCCCTAATAATAGCAAC 1276
Db 1185 AATGGTGTATTTAATTTTCTCCATCTCTCATTAACACAGGGATGCCCTAAGATAGCAAC 1244

Qy 1277 CTTGTCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1333
Db 1245 CC--CATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1302

Qy 1334 ATAGCCAGAGTTAAATAATAATAACAATAATAATAATAATAATAATAATAATAATAATAATA 1393
Db 1303 ATAGCTGAGTTAAATAATAATAACAATAATAATAATAATAATAATAATAATAATAATAATA 1358

Qy 1394 AAAGAATTTCTGGTAAATGATGTGAAATAATGACAGCTCCCTCCTCTCTCTCTCTCTCTCT 1453
Db 1359 AGAAGGCT-----CTGTGAGACTCTTAACAGCTCCTCTCTCTCTCTCTCTCTCTCTCT 1405

Qy 1454 CTATATACAGTCTAGGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1513
Db 1406 CTGTGGAAGGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1465

Qy 1514 AGGCACTGAGAGACATT 1531
Db 1466 AGGCCAGTAAGACAGATT 1483

RESULT 10
LOCUS      AX465671              711 bp          DNA          linear          PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0218425.
ACCESSION  AX465671
VERSION     AX465671.1  GI:21899931
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE   1
  AUTHORS   Meyers, R.A.
  TITLE     47324, a human g-protein and uses therefor
  JOURNAL   Patent: WO 0218425-A 3 07-MAR-2002;
            Millennium Pharmaceuticals (US)
FEATURES    Location/Qualifiers
             1..711
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"

ORIGIN
Query Match      42.0%; Score 711; DB 6; Length 711;
Best Local Similarity 100.0%; Pred. No. 9.9e-164;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 ATGGCGTCCCTGGATCGGTTGAGGTAAGTACTGTTGTTGGAGACTCAGGTTGTTGGAAATCT 78
Db 1 ATGGCGTCCCTGGATCGGTTGAGGTAAGTACTGTTGTTGGAGACTCAGGTTGTTGGAAATCT 60
```


QY 481 ACAAAGCGCATGAAGTTTAACTAGGACTGCTTTCTGCTGAGGATTTCAATCCAGAA 540
 Db 497 GCGAGCGCAATGAGTCTGTCACAGACTGCCCTTCTAGCTGAGGATTTAATGCTGAA 556
 QY 541 GAAATTAATTTGGACTGTCACAAATCCACGGTACTTAGCTGCGAGTCTTCTTCAATGCTGTC 600
 Db 557 GAGATTAATCTGACTGTCAGCAACACACAGATGCTGCTGCTGCTTCTTCAATGCTGTC 616
 QY 601 AAGCTCAGTAGGTTTTTTGATAGGTCATGAGAGAGATACATTTTAAAGAGAGGTAAT 660
 Db 617 AAACCTCAGCAGGTTTTTTTGACAAGGTTATTGAGAAAAGATAC---CCACGGGAAGGAAAT 673
 QY 661 CAGATCCAGCGTTCTCCTGATCGAAAGATTTGGGCGACGACATTTAAAGAGCCTTCAT 720
 Db 674 CTGATCCCTGGTTCTCCGATCGGAACGCTTTGGAGGGGGAGACTTTAAGAGTCTCCAC 733
 QY 721 TATGACTGA 729
 Db 734 TATGACTGA 742

RESULT 13
 BC078191
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BC078191 1040 bp mRNA linear VRT 26-JUL-2004
 Danio rerio cDNA clone MGC:100876 IMAGE:7141370, complete cds.
 BC078191 GI:50603855
 MGC.
 Danio rerio (zebrafish)
 Danio rerio
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 1040)

REFERENCE
 AUTHORS
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Frange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullihy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Vallaloon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1040)
 Strausberg, R.
 Direct Submission
 Submitted (19-JUL-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

TITLE
 JOURNAL
 PUBMED
 REFERENCE

12477932
 2 (bases 1 to 1040)
 Strausberg, R.
 Direct Submission
 Submitted (19-JUL-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Genome Institute of Singapore
 cDNA Library Preparation: S. Mathavan, Chia-Lin Wei, and Yijun
 Ruan, Genome Institute of Singapore
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hri.nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 60 Row: e Column: 9
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Similarity but not identity to protein.

FEATURES

source

1..1040
 Location/Qualifiers
 /organism="Danio rerio"
 /mol_type="mRNA"
 /strain="Singapore local strain"
 /db_xref="taxon:7955"
 /clone="MGC:100876 IMAGE:7141370"
 /tissue_type="Embryo, 7 different stages (from just
 fertilized embryos to 72 hours just hatched baby fish)"
 /clone_lib="GISZP001_ra"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 42..743
 /codon_start=1
 /product="Unknown (protein for MGC:100876)"
 /protein_id="AAH78191.1"
 /db_xref="GI:50603856"
 /translation="MASLDKRVKLVLDGSGVKGKSLVHLHCQVNLGNPSWTVGCSVD
 VRVHDTREGTEPEKAFYIELVDVGVSGVSGVSKVSVFVNSVNGIILVHDI.TNKKK
 SNLVHVSLEALS.KOSSPTGIIVSNGDYDREQFAENAVPLLLGTGFDQI.PENKRNVDV
 LRTFAELSEDFNAEELNCTNPRYLAAGSSNAVKLSRFEDKVIKRYFTRDFSQMS
 FIDRRFRNFSLHSD"

CDS

Query Match 23.4%; Score 397; DB 5; Length 1040;
 Best Local Similarity 74.2%; Pred. No. 1.3e-86;
 Matches 502; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

ORIGIN

QY 17 AGATGGCGTCCCTCGATCGGGTGAAGTACTGTGTGGGAGACTCAGGTGTGGGAAAT 76
 Db 40 AAATGCATCTTTGGACAGGGTGAAGTGTCTGTTTTTGGGAGATTTCTGGAGTAGGGAAGT 99
 QY 77 CTTGCTTAGTCCATCTCCTATGCCAAATCAAGTGTGGGAAATCCATCATGACTGTGG 136
 Db 100 CCTCTCTGTACACCTCTCTTTGCCAGATCAGGTTTTTGGGAAATCCCTCATGACTGTGG 159
 QY 137 GCTGCTCAGTGGATGTGACAGTTCATGATTACAAAGAGAAACCCACAGAGAGAGACCT 196
 Db 160 GCTGCTCAGTGGACGTGAGGTTTCAGACTACAGAGAGGACCTCCAGAGAGAGAGGCTT 219
 QY 197 ACTACATAGAAATTATGGGATTTGGAGGCTCTGTGGGAGTGCAGCAGCAGGTGAAAGCA 256
 Db 220 TCTACATTTGAACCTCTGGGATTTGGAGGATCTGTGGCAGTGCAGCAGGTGTTAAAAAGCA 279
 QY 257 CAGAGCAGTATTCTCAACTCCGTAATGTGATATTTTCGTACACGACTTACAAATA 316
 Db 280 CCAGAGCTGTGTTTTTCAATTTCAGTTAAATGGTATCATTTTAGTTTCAGATCTGACTAACA 339
 QY 317 AGAAGTCTCTCCCAAACTTGGCTGCTGTGGTCAATTTGGAAGCTCTCAACAGGGAATTTGGTGC 376
 Db 340 AGAAGTCTCTCCCAAGATCTATACCGCTGGTCAATTAGAGGCACCTAAGCAGGACTCTCTCTC 399
 QY 377 CAACCTGAGTCTTTGGTGCACAAATGGGGATTTATGATCAGAACACAGTTTGTCTGATACCAA 436
 Db 400 CAACCTGGCATCATCGTATATCCACGGTGTATACGATAGAGAACAGCTTTTGCAGAGAACCGTGT 459

Db 362 TTGGTGCCCAACTGGAGTCTTTGGTG 385

Search completed: June 30, 2005, 23:29:32
Job time : 7774.11 secs

```
CC region 1. .347
CC id AA023764
CC est
CC blastn
CC identity 95
CC region 145. .384
CC id C03036
CC est
CC blastn
CC identity 98
CC region 2. .71
CC id C03036
CC est
CC blastn
CC identity 99
CC region 1. .193
CC id R08519
CC est
CC blastn
CC identity 94
CC region 193. .263
CC id R08519
CC est
CC Von Heijne matrix
CC score 4.8
CC seq SLVHLLCQNQLG/NP
CC n=a, g, c or t
FH Key Location/Qualifiers
FT misc_feature 39. .385
FT misc_feature 146. .385
FT misc_feature 11. .80
FT misc_feature 39. .231
FT misc_feature 232. .302
FT sig_peptide 11. .109
FT misc_feature 287
FT misc_feature 331.
FEATURES
source
1. .385
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 22.1%; Score 374.8; DB 6; Length 385;
Best Local Similarity 97.4%; Pred. No. 4.2e-81;
Matches 374; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 10 AGTGGCAAGATGGCTCCCTGGATCGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTT 69
DB 2 AGTGGCAAGATGGCTCCCTGGATCGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTT 61

QY 70 GGGAAATCTTCGTTAGTCCATCTCCTATGCCAAAATCAAGTCTGGGAAATCCATCATGG 129
DB 62 GGGAAATCTTCGTTAGTCCATCTCCTATGCCAAAATCAAGTCTGGGAAATCCATCATGG 121

QY 130 ACTGTGGGCTGCTCAGTGGATGTCAGATTCATGATTACAAAGAGAACCCAGAGAG 189
DB 122 ACTGTGGGCTGCTCAGTGGATGTCAGATTCATGATTACAAAGAGAACCCAGAGAG 181

QY 190 AAGACCTACTACATAGAAATTTATGGGATGTTGGAGGCTCTGTGGGAGTGCAGCAGCGTG 249
DB 182 AAGACCTACTACATAGAAATTTATGGGATGTTGGAGGCTCTGTGGGAGTGCAGCAGCGTG 241

QY 250 AAAAGCACAAGACAGCATTTCTCAACTCCGTAATGGTATTATTTTCGTACACGACTTA 309
DB 242 AAAAGCACAAGACAGCATTTCTCAACTCCGTAATGGTATTATTTTCGTACACGACTTA 301

QY 310 ACAAAATAGAGTCTCCCAAACTTCGTCGTGTCATTGGAGCTCTCAACAGGAT 369
DB 302 ACVSATGGGAAAGTCTCCCAAAAMTTGCNCGTTGGTCATTGGAGCTCTCAACAGGAT 361

QY 370 TTGGTGCCCAACTGGAGTCTTTGGTG 393
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 13:26:53 ; Search time 1017.81 Seconds
(without alignments)
9852.580 Million cell updates/sec

Title: US-09-945-173-1
Perfect score: 1694
Sequence: 1 aggaagcagtggaagat.....aaacatgaattattgaacta 1694

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004as:*
- 13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1692.4	99.9	1694	6	AAD34049 Human G-p
2	1664.6	98.3	2398	10	AAD54288 Human SEC
3	1358.8	80.2	1417	8	ABX63467 Human CDN
4	735	43.4	735	10	ADAE47763 Human NOV
5	735	43.4	735	12	ADJ79033 Human NOV
6	731.6	43.2	739	10	ADAE47765 Human NOV
7	731.6	43.2	739	12	ADJ79035 Human NOV
8	628.4	37.1	959	4	AAS27286 cDNA enco
9	628.4	37.1	959	10	ACH36304 Human end
10	444	26.2	504	9	ACH36304 Human end
11	374.8	22.1	385	2	AAx41283 Human sec
12	321	18.9	321	3	AAx41283 Human sec
13	252.8	14.9	387	8	ABX52440 Bovine ES
14	66.8	3.9	919	4	ABL08307 Drosophil
15	60	3.5	60	6	ABN37775 Human spl
16	56.8	3.4	3048	4	ABL08306 Drosophil
17	56.8	3.4	3215	4	ABL14880 Drosophil
18	56.8	3.4	3888	4	ABL14882 Drosophil
19	44.2	2.6	1234	3	AAC59835 Human sec
20	44.2	2.6	11922	3	AAA70187 Plasmodiu

21	44	2.6	6713	6	ABL33081 Human imm
22	44	2.6	6713	6	ABL34527 Human met
23	44	2.6	6713	6	ABL70252 Chemicall
24	44	2.6	6713	7	AD999788 Complemen
25	43.8	2.6	751	4	AAI95038 Human neu
26	43.8	2.6	2000	8	ADA71938 Rice gene
27	43.4	2.6	5641	6	ABL33396 Human imm
28	42.4	2.5	444	5	ADL43671 Human ova
29	42	2.5	9789	2	AAAT41852 cDNA enco
30	41.2	2.4	56153	4	AAS46793 Tumour su
31	41.2	2.4	2239	4	AAH55047 S. epider
32	41.2	2.4	2259	4	AAH54714 S. epider
33	41.2	2.4	3475	4	AAH54668 S. epider
34	41.2	2.4	13131	6	ABL92248 Chemicall
35	41.2	2.4	56153	4	AAS46794 Tumour su
36	41	2.4	2000	8	ADA71938 Rice gene
37	40.8	2.4	457	9	ACH49768 Human leu
38	40.8	2.4	495	13	ACN49798 Cotton pr
39	40.8	2.4	176001	12	ADK43203 Human pro
40	40.8	2.4	186739	12	ADK43195 Human pro
41	40.6	2.4	7305	4	AAK90435 Human dig
42	40.6	2.4	7305	5	AAK16679 Human pan
43	40.4	2.4	120239	13	ABD33244 Murine ca
44	40.2	2.4	5461	6	ABL32999 Human imm
45	40.2	2.4	5461	6	ABQ66978 Human ang

ALIGNMENTS

RESULT 1

AAD34049
ID AAD34049 standard; cDNA; 1694 BP.

XX AAD34049;

DT 16-JUL-2002 (first entry)

XX Human G-protein (47324) cDNA.

Human; G-protein; 47324; cancer; signal transduction; tumour; apoptosis; cardiovascular disorder; liver; autoimmune; blood vessel; bone; hormonal; reproductive; fertility; brain; haematopoietic; metabolic; gene therapy; hyperextension; atherosclerosis; Gaucher's disease; degenerative disease; multiple sclerosis; glycogen storage disease; obesity; acute meningitis; anorexia; cerebral oedema; pain; Alzheimer's disease; diabetes mellitus; Parkinson's disease; rheumatoid arthritis; osteoporosis; dermatological; transgenic animal; antiinflammatory; antiinflammatory; vasotrophic; anti-HIV; nootropic; vulnary; ophthalmological; virucide; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT 5'UTR 1..18

FT CDS /*tag= a

FT /*tag= b

FT /*product= "Human 47324 protein"

FT /*note= "CDS referred to as SEQ ID NO:3 is specifically

FT 3'UTR 730..1694

FT /*tag= c

XX WO200218425-A2.

XX PD 07-MAR-2002.

XX 31-AUG-2001; 2001WO-US027197.

XX 01-SEP-2000; 2000US-0229293P.

XX (MILL-) MILLENNIUM PHARM INC.

XX

PI	Meyers RA;
XX	WPI; 2002-315535/35.
DR	P-PSDB; AAE21568.
XX	A human G protein and nucleic acid molecule encoding it, for diagnosing
XX	and treating cellular proliferative, bone, immune, cardiovascular, liver,
PT	pain or metabolic disorders and identifying modulators.
PT	Claim 1; Fig 1; 118pp; English.
XX	The invention relates to a human G-protein, termed 47324 and nucleic acid
CC	encoding it. 47324 is useful for treating cancer or aberrant cellular
CC	proliferation and/or differentiation such as breast, ovarian, prostate,
CC	colon and lung cancer, pain, heart disorders and viral disease. 47324
CC	associate a lipid bilayer due to post-translational modifications,
CC	modulate signal transduction, control cell morphology and
CC	differentiation, mediate cellular proliferation, apoptosis, modulate
CC	tumour inhibition, regulate cellular trafficking and active cytoskeleton
CC	organisation. Cardiovascular disorders include hypertension and
CC	atherosclerosis; liver disorders include Gaucher's disease and glycogen
CC	storage disease; metabolic disorders include obesity and anorexia; brain
CC	disorders include cerebral oedema and neural tube defect; infections
CC	include acute meningitis and multiple sclerosis; degenerative diseases
CC	include Alzheimer's disease and Parkinson's disease; blood vessel
CC	disorders include responses of vascular cell walls to injury;
CC	haematopoietic disorders include autoimmune disease such as diabetes
CC	mellitus and rheumatoid arthritis; bone disorders include osteoporosis;
CC	hormonal disorders include reproductive or fertility disorders. Ras
CC	family-associated or related disorders also include immune disorders such
CC	as autoimmune disorders or immune deficiency disorders e.g. severe
CC	combined immunodeficiency. 47324 is also useful as pharmacogenomic
CC	markers. 47324 is useful for producing non-human transgenic animals which
CC	are useful for studying the function and/or activity of 47324 and for
CC	identifying and/or evaluating modulators of 47324 activity. 47324 gene is
CC	useful in gene therapy. The present sequence is human 47324 cDNA
XX	Sequence 1694 BP; 530 A; 323 C; 360 G; 481 T; 0 U; 0 Other;
SQ	Query Match 99.9%; Score 1692.4; DB 6; Length 1694; Best Local Similarity 99.9%; Pred. No. 0; Matches 1693; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 AGGGAAGCAGTGGCAAGATGGCTCCGTGATCGGGTGAAAGTACTCGTTGGGAGAC 60
DB	1 AGGGAAGCAGTGGCAAGATGGCTCCGTGATCGGGTGAAAGTACTCGTTGGGAGAC 60
QY	61 TCAGGTGTTGGGAAATCTTCGTTAGTCCATCTCTATGCCAAAATCAAAGTCTCGGAAAT 120
DB	61 TCAGGTGTTGGGAAATCTTCGTTAGTCCATCTCTATGCCAAAATCAAAGTCTCGGAAAT 120
QY	121 CCATCATGAGACTGTGGGCTGCTCAGTGGATGCAGATTTCATGATCAAGAAGGAACC 180
DB	121 CCATCATGAGACTGTGGGCTGCTCAGTGGATGCAGATTTCATGATCAAGAAGGAACC 180
QY	181 CCAGAAGAGAACCTACTACATAGAAATTAATGGATGTTGGAGGCTCTGTGGCAGTGCC 240
DB	181 CCAGAAGAGAACCTACTACATAGAAATTAATGGATGTTGGAGGCTCTGTGGCAGTGCC 240
QY	241 AGCAGCGTGAAGAAGCAACAAGCAGTATTTCTAACAATCCGTAATSGTATTTATTTTCGTA 300
DB	241 AGCAGCGTGAAGAAGCAACAAGCAGTATTTCTAACAATCCGTAATSGTATTTATTTTCGTA 300
QY	301 CAGCATTAACAAATAAGAAAGTCTCCAAAACTTGGCGTGTGGTCAATGGAAAGCTCTC 360
DB	301 CAGCATTAACAAATAAGAAAGTCTCCAAAACTTGGCGTGTGGTCAATGGAAAGCTCTC 360
QY	361 AACAGGGATTTGGTGCACACTGGAGCTTGGTGACAAATGGGGATTAATGATCAAGAACAG 420
DB	361 AACAGGGATTTGGTGCACACTGGAGCTTGGTGACAAATGGGGATTAATGATCAAGAACAG 420
QY	421 TTTCGTGATAACCAATATACCACCTGTTGGTAAATAGGGAATCAAACTGGACCAGATTCATGAA 480
DB	421 TTTCGTGATAACCAATATACCACCTGTTGGTAAATAGGGAATCAAACTGGACCAGATTCATGAA 480
DB	421 TTTCGTGATAACCAATATACCACCTGTTGGTAAATAGGGAATCAAACTGGACCAGATTCATGAA 480
QY	481 ACAAGCGCCATGAAGATTTTTAACTAGGACTGCTTTTCTCGGCTGAGGATTTTCAATCCAGAA 540
DB	481 ACAAGCGCCATGAAGATTTTTAACTAGGACTGCTTTTCTCGGCTGAGGATTTTCAATCCAGAA 540
QY	541 GAAATTAATTTGGACTGCACAAAATCCAAGTACTTAGCTGAGGATTTTCTTCAATGCTGCTC 600
DB	541 GAAATTAATTTGGACTGCACAAAATCCAAGTACTTAGCTGAGGATTTTCTTCAATGCTGCTC 600
QY	601 FAGCTCAGTAGGATTTTTGATAAGTCTATAGAGAGAGATCTTTTTTAAAGRAGAGGTAAT 660
DB	601 FAGCTCAGTAGGATTTTTGATAAGTCTATAGAGAGAGATCTTTTTTAAAGRAGAGGTAAT 660
QY	661 CAGATTCCAGGCTTTTCTGATCGGAAAAGATTTGGGGCAGAGAAATTTAAAGAGCCTTCAT 720
DB	661 CAGATTCCAGGCTTTTCTGATCGGAAAAGATTTGGGGCAGAGAAATTTAAAGAGCCTTCAT 720
QY	721 TAGTACTGAATPACACTCATCTCTTTTGGAAAGATGAGCAAGCAGTGGCGAGTTTTCACAGC 780
DB	721 TAGTACTGAATPACACTCATCTCTTTTGGAAAGATGAGCAAGCAGTGGCGAGTTTTCACAGC 780
QY	781 TCATCTTGTCTGTGTTCAATATTATACCATCACAGCCTTTTAAACAAATCATCTTAAATGTC 840
DB	781 TCATCTTGTCTGTGTTCAATATTATACCATCACAGCCTTTTAAACAAATCATCTTAAATGTC 840
QY	841 TACCTTTCAGCCTTACCTTTAAATCGAAAAATGAAGGAAGTGAACAATACGGGAGGTCCA 900
DB	841 TACCTTTCAGCCTTACCTTTTANTGAAAAATGAAGGAAGTGAACAATACGGGAGGTCCA 900
QY	901 AACTTTTGTCCCTGTTCTCTGTGTTCTTACCTTTCTGTCCTCTGTGTATAGATTAATGTA 960
DB	901 AACTTTTGTCCCTGTTCTCTGTGTTCTTACCTTTCTGTCCTCTGTGTATAGATTAATGTA 960
QY	961 AGCTTTGTGTAATATAGATGTTGTCAAAATGATGCAAGTAAATGAGCAATGACAGTGTA 1020
DB	961 AGCTTTGTGTAATATAGATGTTGTCAAAATGATGCAAGTAAATGAGCAATGACAGTGTA 1020
QY	1021 CTGCAGAGAAAATTTACTCTTCCTCCTAGAACCTGGAGGGTTTTTATGGGGTCTGTAATTTTC 1080
DB	1021 CTGCAGAGAAAATTTACTCTTCCTCCTAGAACCTGGAGGGTTTTTATGGGGTCTGTAATTTTC 1080
QY	1081 CACACTCATTTGCTGAAAGCTTAAATTAAGTACTTCAAAAACGTAATCTCCATTTGTTTACCT 1140
DB	1081 CACACTCATTTGCTGAAAGCTTAAATTAAGTACTTCAAAAACGTAATCTCCATTTGTTTACCT 1140
QY	1141 TCTTGAGGGAAACGGTCTTTGTTAAACAGCCCTGAGTTGTCTACCCCAAAACAATCTCTGTC 1200
DB	1141 TCTTGAGGGAAACGGTCTTTGTTAAACAGCCCTGAGTTGTCTACCCCAAAACAATCTCTGTC 1200
QY	1201 ATTTTCAAGATGCAAAATGGTGTATTAATTTGTTCTCCACATTTGTCTCACACACAGGAAT 1260
DB	1201 ATTTTCAAGATGCAAAATGGTGTATTAATTTGTTCTCCACATTTGTCTCACACACAGGAAT 1260
QY	1261 GCCTAATAATGCAAAACCTTGTCTCCCTCTCTCTCTCTTTTGC AAAATGGCTCAGTGACTGG 1320
DB	1261 GCCTAATAATGCAAAACCTTGTCTCCCTCTCTCTCTCTCTTTTGC AAAATGGCTCAGTGACTGG 1320
QY	1321 AAGAGCGGACTAATAGCCAGAGTTAAATATAATAATAAATAAATAAATAAATAAATAAATAA 1380
DB	1321 AAGAGCGGACTAATAGCCAGAGTTAAATATAATAATAAATAAATAAATAAATAAATAAATAA 1380
QY	1381 GCNATACAGAAAAAAGNAATTCGTGTAATATGATGTGAAAAATTTGACAGCTCCCTCACT 1440
DB	1381 GCNATACAGAAAAAAGNAATTCGTGTAATATGATGTGAAAAATTTGACAGCTCCCTCACT 1440
QY	1441 CTTAAGTGTGCTGTATATACAGTCTAGGTTTTCTGTTTGGAAATAGGTAGGGTAAAAATC 1500
DB	1441 CTTAAGTGTGCTGTGTATATACAGTCTAGGTTTTCTGTTTGGAAATAGGTAGGGTAAAAATC 1500
QY	1501 TAAGACCTGCAACAAGGCGAGTGAAGACATTTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCT 1560
DB	1501 TAAGACCTGCAACAAGGCGAGTGAAGACATTTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCT 1560

QY 1561 GGAAGTCAACTCTCTGAAATGTCCCTTAGCTATAATCAGAAACTAAGAAATATTCT 1620
 |||||
 Db 1561 GGAAGTCAACTCTCTGAAATGTCCCTTAGCTATAATCAGAAACTAAGAAATATTCT 1620
 |||||
 QY 1621 GTGTCAACATGTATTATTGGAGAGAGTAAATTAAGTTCCACAGCAACAAAAACAT 1680
 |||||
 Db 1621 GTGTCAACATGTATTATTGGAGAGAGTAAATTAAGTTCCACAGCAACAAAAACAT 1680
 |||||
 QY 1681 GAATATTGAACTA 1694
 |||||
 Db 1681 GAATATTGAACTA 1694
 |||||

RESULT 2

AAD54288
 ID AAD54288 standard; cDNA; 2398 BP.

AC AAD54288;

XX 17-JUN-2003 (first entry)

XX Human SECP-2 cDNA.

XX Human; secreted protein; SECP; cell proliferative disorder; cirrhosis;
 KW cancer; arteriosclerosis; hepatitis; psoriasis; developmental disorder;
 KW atherosclerosis; seizure disorder; renal tubular acidosis; anaemia;
 KW cataract; sensorineural hearing loss; neurological disorder; epilepsy;
 KW ischaemic cerebrovascular disease; Alzheimer's disease; Pick's disease;
 KW Huntington's disease; stroke; dementia; Parkinson's disease; anxiety;
 KW amyotrophic lateral sclerosis; schizophrenia disorder; mental disorder;
 KW Tourette's disorder; muscular dystrophy; autoimmune disorder; mood; AIDS;
 KW inflammatory disorder; acquired immunodeficiency syndrome; allergy; ARDS;
 KW adult respiratory distress syndrome; diabetes mellitus; Crohn's disease;
 KW asthma; autoimmune thyroiditis; glomerulonephritis; rheumatoid arthritis;
 KW atopic dermatitis; ulcerative colitis; trauma; cardiovascular disorder;
 KW infection; congestive heart failure; heart disease; angina pectoris;
 KW myocardial infarction; myocarditis; transgenic animal; gene therapy;
 KW transgenic; vaccine; virucide; antibacterial; fungicide; antiparasitic;
 KW protozoacide; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

CDS 3..713
 /*tag= a
 /product= "Human SECP-2 protein"

FT sig_peptide 3..101

FT mat_peptide 102..710

FT /*tag= b

FT /*tag= c

FT /product= "Human mature SECP-2 protein"

XX W0200297035-A2.

XX 05-DEC-2002.

XX 21-MAY-2002; 2002MO-US016234.

XX 25-MAY-2001; 2001US-0293728P.

XX 08-JUN-2001; 2001US-0297019P.

XX 19-JUN-2001; 2001US-0292979P.

XX 22-JUN-2001; 2001US-0300537P.

XX 29-JUN-2001; 2001US-0301936P.

XX 06-MAR-2002; 2002US-0362439P.

XX 19-MAR-2002; 2002US-0366041P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Lee EA, Becha SD, Baughn MR, Yao MG, Tang YT;

XX Au-Young JK, Lal PG, Warren BA, Duggan BM, Tran UK, Xu Y;

XX Thangavelu K, Richardson TW, Bandman O, Jones KA, Yang J;

PI Emerling BM, Swarnakar A, Luo W, Walia NK, Azimzai Y, Khan FA;
 PI Lu DM, Griffin JA, Lee SY, Burford N, Elliott VS, Honcheil CD;
 PI He A, Mason PM, Li JX, Hafalia AJA, Gururajan R;
 XX WPI; 2003-129519/12.
 DR P-PSDB; AAE35750.

XX Novel human secreted proteins and genes encoding the proteins, useful for
 PT treating, diagnosing and preventing cell proliferative,
 PT autoimmune/inflammatory, cardiovascular, developmental or neurological
 PT disorders.

PS Claim 89; Page 171-172; 192pp; English.

XX The present invention relates to novel human secreted proteins (SECP) and
 CC polynucleotides encoding such proteins. SECP sequences are useful for
 CC diagnosing, treating and preventing cell proliferative disorders
 CC including cancer (e.g. arteriosclerosis, cirrhosis, hepatitis, psoriasis
 CC and atherosclerosis), developmental disorders (e.g. seizure disorders,
 CC renal tubular acidosis, anaemia, cataract and sensorineural hearing
 CC loss), neurological disorders (e.g. epilepsy, ischaemic cerebrovascular
 CC disease, Alzheimer's disease, Pick's disease, Huntington's disease,
 CC stroke, dementia, Parkinson's disease, amyotrophic lateral sclerosis,
 CC schizophrenic disorders, mental disorders including mood and anxiety,
 CC Tourette's disorder and muscular dystrophy), autoimmune/inflammatory
 CC disorders (e.g. acquired immunodeficiency syndrome (AIDS), allergy, adult
 CC respiratory distress syndrome (ARDS), asthma, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, glomerulonephritis, rheumatoid
 CC arthritis, atopic dermatitis, ulcerative colitis, trauma and viral
 CC bacterial, fungal, parasitic, protozoal and helminthic infections) and
 CC cardiovascular disorders (e.g. congestive heart failure, ischaemic heart
 CC disease, angina pectoris, myocardial infarction, hypertensive heart
 CC disease, congenital heart disease and myocarditis). They are useful for
 CC creating knocking humanised animals or transgenic animals to model human
 CC diseases. Sequences of the invention are useful in somatic or germline
 CC gene therapy and in diagnostic purposes. They are also used as vaccines.
 CC The present sequence is human SECP-2 cDNA

XX Sequence 2398 BP; 722 A; 456 C; 470 G; 750 T; 0 U; 0 Other;

Query Match 98.3%; Score 1664.6; DB 10; Length 2398;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1667; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 17 AGATGGCGTCCCTGGATCGGGTCAAGGTACTGTGTGGGAGACTCAGGTGTGGGAAT 76

Db 1 AGATGGCGTCCCTGGATCGGGTCAAGGTACTGTGTGGGAGACTCAGGTGTGGGAAT 60

QY 77 CTTGGTTAGTCCATCTCTATGCCAAATCAAGTGTGGGAAATCCATCATGGACTGTGG 136

Db 61 CTTGGTTAGTCCATCTCTATGCCAAATCAAGTGTGGGAAATCCATCATGGACTGTGG 120

QY 137 GCTGCTCAGTGGAGTGTCCAGAGTTCATGATTACAAAGAGGAACCCAGAGAGAGACCT 196

Db 121 GCTGCTCAGTGGAGTGTCCAGAGTTCATGATTACAAAGAGGAACCCAGAGAGAGACCT 180

QY 197 ACTACATAGATTAATATGGATGTTGGAGGCTCTGTGGCAGTGCAGCAGCGTGAAGAACA 256

Db 181 ACTACATAGATTAATATGGATGTTGGAGGCTCTGTGGCAGTGCAGCAGCGTGAAGAACA 240

QY 257 CAAGAGCAGTATTCTCAACATCCGTAATATGTTATTTTCGTACAGCATTAACAATA 316

Db 241 CAAGAGCAGTATTCTCAACATCCGTAATATGTTATTTTCGTACAGCATTAACAATA 300

QY 317 AGAAGTCTCCCAAACTTGGTGTTCATTTGGAGCTCTCAACAGGGATTTGGTGC 376

Db 301 AGAAGTCTCCCAAACTTGGTGTTCATTTGGAGCTCTCAACAGGGATTTGGTGC 360

QY 377 CAACTGGAGTCTTGGTGACAAAATGGGATTAATGATCAAGAACAGTTTCTCATTAACCAA 436

Db 361 CAACTGGAGTCTTGGTGACAAAATGGGATTAATGATCAAGAACAGTTTCTCATTAACCAA 420

QY 437 TACCACCTGTTGTAATAGGGACTAAACTGGACCAAGATTCATGAACAACAGCCCATGAAG 496

Db 421 TACCAGTGTGGTAAATAGGAGCTAAACTGACCCAGATTCATGAAAACAGAGCGCCATGAAG 480
QY 497 TTTTAACTAGGACTGCTTTCTCGGCTGAGGATTTCAATCAGAGAAAATTAATTTGGACT 556
Db 481 TTTTAACTAGGACTGCTTTCTCGGCTGAGGATTTCAATCAGAGAAAATTAATTTGGACT 540
QY 557 GCACAAATCCACGGTACTTAGCTGCGAGGTTCTTCCAAATGCTGTCAGCTCAGTAGGTTT 616
Db 541 GCACAAATCCACGGTACTTAGCTGCGAGGTTCTTCCAAATGCTGTCAGCTCAGTAGGTTT 600
QY 617 TTGATAAGGTCATAGAGAGAGATACCTTTTAAAGAGAGGTAATCAGATTTCCAGGCTTTTC 676
Db 601 TTGATAAGGTCATAGAGAGAGATACCTTTTAAAGAGAGGTAATCAGATTTCCAGGCTTTTC 660
QY 677 CTGATCGGAAAAGATTTGGGGCAGGAACATTAAGAGAGCCCTTCAATATGACTGAATTAAC 736
Db 661 CTGATCGGAAAAGATTTGGGGCAGGAACATTAAGAGAGCCCTTCAATATGACTGAATTAAC 720
QY 737 TCATCCTTTGGAGAGTGACGACGAGTGGCAGTGTTCACAGCTCATCTTGCTGTGTTTC 796
Db 721 TCATCCTTTGGAGAGTGACGACGAGTGGCAGTGTTCACAGCTCATCTTGCTGTGTTTC 780
QY 797 AATTATTACCATCACAGCCCTTTTAAACAAAATCATCTTAAATGCTACCCCTTCAGCCCTTAC 856
Db 781 AATTATTACCATCACAGCCCTTTTAAACAAAATCATCTTAAATGCTACCCCTTCAGCCCTTAC 840
QY 857 CCTTTAATGGAAAATGAAAGGAGTGACAAATACGGGAGGTCCAAAATTTGTCCCTGTTC 916
Db 841 CCTTTAATGGAAAATGAAAGGAGTGACAAATACGGGAGGTCCAAAATTTGTCCCTGTTC 900
QY 917 TCTGTGTCCTTACCTTTCTGTCCTGCTGTATAGATTTATGTAAGCCCTTGTTGTAATAT 976
Db 901 TCTGTGTCCTTACCTTTCTGTCCTGCTGTATAGATTTATGTAAGCCCTTGTTGTAATAT 960
QY 977 GAGATGTTGTCAAAATGATGCAGTAAATGAGCAATGACAGTGTACTGCAGAGAAAATTTA 1036
Db 961 GAGATGTTGTCAAAATGATGCAGTAAATGAGCAATGACAGTGTACTGCAGAGAAAATTTA 1020
QY 1037 CTCTTGCTAGAACTGGAGGGTTTTATGCGGTCTGTAATTTTCCACACTCATTTGCTGAA 1096
Db 1021 CTCTTGCTAGAACTGGAGGGTTTTATGCGGTCTGTAATTTTCCACACTCATTTGCTGAA 1080
QY 1097 AGCTTAATTAAGTACTTCAAAAACGTTATCTCCATTTGTTTACCTTTTGAGGGGACGGT 1156
Db 1081 AGCTTAATTAAGTACTTCAAAAACGTTATCTCCATTTGTTTACCTTTTGAGGGGACGGT 1140
QY 1157 CTTGTTAAACGAGCCCTGAGTGTCTACCCCAAAACAATCTCTGTCAATTTTCAAAAGATGCAA 1216
Db 1141 CTTGTTAAACGAGCCCTGAGTGTCTACCCCAAAACAATCTCTGTCAATTTTCAAAAGATGCAA 1200
QY 1217 AATGGTGTATTAATTTGCTCCACCAATGTCACACAGGAATGCTTAATAATAGCAAC 1276
Db 1201 AATGGTGTATTAATTTGCTCCACCAATGTCACACAGGAATGCTTAATAATAGCAAC 1260
QY 1277 CTTGTCTCCCTCTCTCTCTTGCAGAAATGGCTGAGTGTGAGAGGCGGCACTAATA 1336
Db 1261 CTTGTCTCCCTCTCTCTCTTGCAGAAATGGCTGAGTGTGAGAGGCGGCACTAATA 1320
QY 1337 GCCAGAGTTAAATATATAAATACAAATTAATAATACATAGAGAACAGCAATACAGAGAAAAA 1396
Db 1321 GCCAGAGTTAAATATATAAATACAAATTAATAATACATAGAGAACAGCAATACAGAGAAAAA 1380
QY 1397 AGAATTTCTGTAATAATGATGAAAAATTCACAGCTCCCTCACTTCAAGGTTGCTGCTA 1456
Db 1381 AGAATTTCTGTAATAATGATGAAAAATTCACAGCTCCCTCACTTCAAGGTTGCTGCTA 1440
QY 1457 TATACAGTCTAGCTTTCTGTTGGAAATAGGTAGGTAAAAATTAAGACCTGCAAGG 1516
Db 1441 TATACAGTCTAGCTTTCTGTTGGAAATAGGTAGGTAAAAATTAAGACCTGCAAGG 1500
QY 1517 GCAGTGAGAGACATTTTACAGCCCTCTCTATTTGTTTTTTTAAAGGAAAAGTCAACTCCT 1576

Db 1501 GCAGTGAGAGACATTTTACAGCCCTCTCTCTATTTTGTGTTTTTAAAGGAAAAGTCAACTCCT 1560
QY 1577 GAAATGTCCCTTAGCTATATATCAGAAAACTCAGAAATATTTATTTCTGTGTCAACAATGATT 1636
Db 1561 GAAATGTCCCTTAGCTATATATCAGAAAACTCAGAAATATTTATTTCTGTGTCAACAATGATT 1620
QY 1637 TATGGAGAGAAAGTAAAAATAAGTTTCCACAGCAACACAAAAACATGAATTAT 1687
Db 1621 TATGGAGAGAAAGTAAAAATAAGTTTCCACAGCAACACATTTACATGAATTAT 1671
RESULT 3
ABX63467
ID ABX63467 standard; cDNA; 1417 BP.
XX AC ABX63467;
XX XX
DT 26-FEB-2003 (first entry)
XX Human cDNA #467 differentially expressed in activated vascular tissue.
DE Human; gene; ss; vascular tissue; cytostatic; atherosclerosis; cardiac;
KW hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;
KW gene therapy; vascular disease; cancer; coronary; artery disease;
KW hypertension; diabetes; pre-eclampsia; resection;
KW ischaemia-reperfusion injury; stroke.
XX OS Homo sapiens.
XX PN US2002137081-A1.
XX PD 26-SEP-2002.
XX PF 08-JAN-2002; 2002US-00044090.
XX PR 28-JUL-2000; 2000US-0222469P.
XX PR 08-JAN-2001; 2001US-0260483P.
XX PA (BAND/) BANDMAN O.
XX PI Bandman O;
XX WPI; 2003-110597/10.
DR Combination for diagnosing, staging, treating, or monitoring the
XX progression of treatment of a vascular disease, e.g. atherosclerosis,
PT comprises several cDNAs that are differentially expressed in activated
PT vascular tissue.
XX Claim 1; Page; 18pp; English.
XX This invention relates to a combination comprising several cDNAs that are
CC differentially expressed in activated vascular tissue. The invention also
CC discloses a high throughput method for detecting differentially expressed
CC cDNAs in a sample. The cDNAs of the invention may have
CC antiatherosclerotic; cytostatic; cardiac; hypotensive; antidiabetic;
CC gynaecological; vasotropic and cerebroprotective activities and may be
CC used in gene therapy. The cDNAs of the invention may be used in a high-
CC throughput methods for detecting differential expression of one or more
CC cDNAs in a sample, or screening several molecules or compounds to
CC identify a molecule or compound that specifically binds a cDNA of the
CC invention. A protein encoded by the cDNA may be used to screen several
CC molecules or compounds to identify a ligand that specifically binds to
CC the protein, or to produce or purify an antibody to the protein that can
CC be used to detect a protein in a sample or purify a natural or
CC recombinant protein from a sample. The nucleotides may be useful for
CC diagnosing, staging, treating, or monitoring the progression of treatment
CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery
CC disease, hypertension, diabetes, pre-eclampsia, ischaemia-reperfusion
CC injury, resection, or stroke. The cDNAs can also be used for large-scale
CC genetic or gene expression analysis of several new nucleic acid
CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for
CC diagnosing pre-pathologic disorders, and chronic or acute diseases

CC associated with abnormalities in the expression, amount or distribution
 CC of the protein. The present sequence represents a cDNA of the invention
 CC that is differentially expressed in activated vascular tissue. Note: The
 CC sequence data for this patent did not form part of the specification, but
 CC was obtained in electronic format directly from USPTO at
 CC <http://seqdata.uspto.gov/sequence.html?DocID=20020137081>
 XX
 SQ Sequence 1417 BP; 426 A; 279 C; 308 G; 404 T; 0 U; 0 Other;

Query Match 80.2%; Score 1358.8; DB 8; Length 1417;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1360; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	17	AGATGGCGTCCCTGGATCGGCTGAAGTACTGGTGTGGGAGACTCAGGTGTTGGGAAAT	76
Db	1	AGATGGCGTCCCTGGATCGGCTGAAGTACTGGTGTGGGAGACTCAGGTGTTGGGAAAT	60
Qy	77	CTTCGGTGTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGGACTGTGG	136
Db	61	CTTCGGTGTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGGACTGTGG	120
Qy	137	GCTGCTCAGTGGATGTGAGAGTTCATGATTAACAAGAGGAAACCCAGAGAGAGACCT	196
Db	121	GCTGCTCAGTGGATGTGAGAGTTCATGATTAACAAGAGGAAACCCAGAGAGAGACCT	180
Qy	197	ACTACATAGAAATTAATGGATGTTGGAGGCTCTGTGGGAGTGCAGCAGCGTGAAGCA	256
Db	181	ACTACATAGAAATTAATGGATGTTGGAGGCTCTGTGGGAGTGCAGCAGCGTGAAGCA	240
Qy	257	CAAGCAGATGTTCTCAACCTCCGTAATGATTAATTTGCTACAGCTTCAACAATA	316
Db	241	CAAGCAGATGTTCTCAACCTCCGTAATGATTAATTTGCTACAGCTTCAACAATA	300
Qy	317	AGAGTCTCTCCAAAACCTTGGCTGTGGTCAATGGAGCTCTCAACAGGGAATTTGGTGC	376
Db	301	AGAGTCTCTCCAAAACCTTGGCTGTGGTCAATGGAGCTCTCAACAGGGAATTTGGTGC	360
Qy	377	CAACTGGAGTCTTGGTGACAAATGGGATTAATGATCAAGAACAGTTGCTGATTAACCAA	436
Db	361	CAACTGGAGTCTTGGTGACAAATGGGATTAATGATCAAGAACAGTTGCTGATTAACCAA	420
Qy	437	TACCACCTGTTGATAGGAGTAACTGGACCAAGTTCATGAACCAAGCGCATGAAG	496
Db	421	TACCACCTGTTGATAGGAGTAACTGGACCAAGTTCATGAACCAAGCGCATGAAG	480
Qy	497	TTTTAACTAGGACTGCTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTTGACT	556
Db	481	TTTTAACTAGGACTGCTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTTGACT	540
Qy	557	GCACAAATCCAGGTAATTTAGTCTGAGGTTCTTCCAAATGCTGCAAGCTCAGTAGGTTT	616
Db	541	GCACAAATCCAGGTAATTTAGTCTGAGGTTCTTCCAAATGCTGCAAGCTCAGTAGGTTT	600
Qy	617	TTGATAGGTCATAGAGAGATGATCTTTTAAAGAGAGTAAATCAGATTCAGGCTTTC	676
Db	601	TTGATAGGTCATAGAGAGATGATCTTTTAAAGAGAGTAAATCAGATTCAGGCTTTC	660
Qy	677	CTGATCGGAAAGATTTGGGCGAGCAATTTAAAGAGCTTTCATTTAGTACGTAATTAAC	736
Db	661	CTGATCGGAAAGATTTGGGCGAGCAATTTAAAGAGCTTTCATTTAGTACGTAATTAAC	720
Qy	737	TCATCTCTTTGGAAGAGTGAAGCAGTGGCAGTGTTCACAGCTCATCTTGTGCTGTC	796
Db	721	TCATCTCTTTGGAAGAGTGAAGCAGTGGCAGTGTTCACAGCTCATCTTGTGCTGTC	780
Qy	797	AATTAATACCATCAGCCTTTTAAACAAATCATCTTAAATGCTACCCCTCAGCCTTAC	856
Db	781	AATTAATACCATCAGCCTTTTAAACAAATCATCTTAAATGCTACCCCTCAGCCTTAC	840
Qy	857	CCTTTAATGGAAATTAAGAGAGTGAACATACGGGAGGTCCTTCTGCTGCTGTC	916
Db	841	CCTTTAATGGAAATTAAGAGAGTGAACATACGGGAGGTCCTTCTGCTGCTGTC	900

Qy	917	TCGTGTTCCTTACCTTCTGTCCCTGTGTATAGATTATGTAAAAGCCTTGTGTAATAT	976
Db	901	TCGTGTTCCTTACCTTCTGTCCCTGTGTATAGATTATGTAAAAGCCTTGTGTAATAT	960
Qy	977	GAGATGTTGTCAAAATGATGACGTAATGAGCAATGACAGTGTACTGCAGAGAAAATTTA	1036
Db	961	GAGATGTTGTCAAAATGATGACGTAATGAGCAATGACAGTGTACTGCAGAGAAAATTTA	1020
Qy	1037	CTCTTGCTAGAACTGAGGGTTTTTATGGGTCTGTAAATTTTCCACACTCATTTGCTGAA	1096
Db	1021	CTCTTGCTAGAACTGAGGGTTTTTATGGGTCTGTAAATTTTCCACACTCATTTGCTGAA	1080
Qy	1097	AGCTTAATTAAGTACTTCAAAAACGTAATCTCCATTTGTACCTTCTTGAGGGGAACCGT	1156
Db	1081	AGCTTAATTAAGTACTTCAAAAACGTAATCTCCATTTGTACCTTCTTGAGGGGAACCGT	1140
Qy	1157	CTTGTAACACAGCCCTGAGTTGTCTAGCCCAAAACAAATCTCTGTCAATTTTCAAGATGCA	1216
Db	1141	CTTGTAACACAGCCCTGAGTTGTCTAGCCCAAAACAAATCTCTGTCAATTTTCAAGATGCA	1200
Qy	1217	AATGCTGTTAATTAATTTGTTCTCCACCATTTGTCACACAGGAATGCCTTAATAGCAAC	1276
Db	1201	AATGCTGTTAATTAATTTGTTCTCCACCATTTGTCACACAGGAATGCCTTAATAGCAAC	1260
Qy	1277	CCTTGCTCCCTCTTCTCTCTCTTTGCAATGGCTCAGTGAAGGGCGGACTAATA	1336
Db	1261	CCTTGCTCCCTCTTCTCTCTCTTTGCAATGGCTCAGTGAAGGGCGGACTAATA	1320
Qy	1337	GCCAGAGTTAAATAATAACAAATTAATAATACATAGAGAA	1378
Db	1321	GCCAGAGTTAAATAATAACAAATTAATAATACATAGAGAA	1362

RESULT 4

AD	ADE47763	standard; DNA; 735 BP.
XX	AD	ADE47763
AC	AD	ADE47763
XX	AD	ADE47763
DT	29-JAN-2004	(first entry)
XX	Human NOV38a	gene SEQ ID NO:125.
DE	ds	gene; human; cardiant; antiarteriosclerotic; hypotensive; immunosuppressive; dermatological; anorectic; cytostatic; antidiabetic; haemostatic; anti-HIV; antiasthmatic; antibacterial; virucide; neuroprotective; nootropic; antiparkinsonian; antilipemic; gene therapy; vaccine.
XX	Homo sapiens.	
OS	WO2003076642-A2.	
XX	WO2003076642-A2.	
XX	18-SEP-2003.	
XX	02-AUG-2002;	2002WO-US024459.
PF	02-AUG-2001;	2001US-0309501P.
XX	03-AUG-2001;	2001US-0310291P.
PR	08-AUG-2001;	2001US-0310951P.
PR	09-AUG-2001;	2001US-0311292P.
PR	13-AUG-2001;	2001US-0311979P.
PR	14-AUG-2001;	2001US-0312203P.
PR	17-AUG-2001;	2001US-0313156P.
PR	17-AUG-2001;	2001US-0313201P.
PR	20-AUG-2001;	2001US-0313702P.
PR	21-AUG-2001;	2001US-0314031P.
PR	23-AUG-2001;	2001US-0314466P.
PR	28-AUG-2001;	2001US-0315403P.
PR	29-AUG-2001;	2001US-0315853P.
PR	31-AUG-2001;	2001US-0316508P.
PR	21-SEP-2001;	2001US-0323936P.
PR	03-DEC-2001;	2001US-0338078P.

20-AUG-2001; 2001US-0313702P.
 21-AUG-2001; 2001US-0314031P.
 23-AUG-2001; 2001US-0314466P.
 28-AUG-2001; 2001US-0315403P.
 29-AUG-2001; 2001US-0315853P.
 31-AUG-2001; 2001US-0316508P.
 17-SEP-2001; 2001US-0322716P.
 21-SEP-2001; 2001US-0323936P.
 03-DEC-2001; 2001US-0338078P.
 05-FEB-2002; 2002US-0354655P.
 05-MAR-2002; 2002US-0361764P.
 19-APR-2002; 2002US-0373825P.
 15-MAY-2002; 2002US-0380971P.
 15-MAY-2002; 2002US-0380980P.
 16-MAY-2002; 2002US-0381039P.
 28-MAY-2002; 2002US-0383761P.
 29-MAY-2002; 2002US-0383887P.
 (ZERH/) ZERHUSEN B D.
 (PATT/) PATTURAJAN M.
 (KEKU/) KEKUDA R.
 (MILL/) MILLER C E.
 (RIEG/) RIEGER D K.
 (PENA/) PENA C E A.
 (SHIM/) SHIMKETS R A.
 (LILL/) LI L.
 (BERG/) BERGHS C.
 (ZHON/) ZHONG M.
 (CASM/) CASMAN S J.
 (VOSS/) VOSS E Z.
 (BOLD/) BOLDIG F L.
 (PADI/) PADIGARU M.
 (SMIT/) SMITHSON G.
 (JIW/) JI W.
 (GORM/) GORMAN L.
 (VERN/) VERNET C A M.
 (LEIT/) LEITE M W.
 (GUOX/) GUO X S.
 (ANDE/) ANDERSON D W.
 (SPYT/) SPYTEK K A.
 (GERL/) GERLACH V.
 (BURG/) BURGESS C E.
 (KHRA/) KHRAMTSOV N V.
 (ORTT/) ORT T.
 (ELLE/) ELLERMAN K.
 (RAST/) RASTELLI L.
 (AGEE/) AGEE M L.
 (CHAU/) CHAUDHURI A.
 (DIPI/) DIPIPPO V A.
 (EDIN/) EDINGER S R.
 (EISE/) EISEN A J.
 (GANG/) GANGOLLI E A.
 (GIOT/) GIOT L.
 (OOIC/) OOI C E.
 (ROTH/) ROTHENBERG M E.
 (SPAD/) SPADERNA S K.
 (HJAL/) HJALT T.
 (LIUX/) LIU X.
 (TAUP/) TAUPIER R J.
 (CATT/) CATTERTON E.
 (SHEN/) SHENOY S G.
 ZERHUSEN BD, Patturajan M, Kekuda R, Miller CE, Rieger DK;
 Pena CE, Shimkets RA, Li L, Berghs C, Zhong M, Casman SJ, Voss EZ;
 Boldog FL, Padigar M, Smithson G, Ji W, Gorman L, Vernet CM;
 Leite MW, Guo XS, Anderson DW, Spytek KA, Gerlach V, Burgess CE;
 Khrantsov NV, Ort T, Ellerman K, Rastelli L, Agee ML, Chaudhuri A;
 Chant JS, Dipippo VA, Edinger SR, Eisen AJ, Gangolli EA, Giot L;
 Ooi CE, Rothenberg ME, Spaderna SK, Hjalt T, Liu X, Taupier RJ;
 Catterton E, Shenoy SG;
 WPI; 2004-108206/11.

DR P-PSDB; ADJ79034.
 XX New isolated NOVX polypeptides and nucleic acid molecules useful for
 PT treating, preventing and diagnosing pathological conditions with NOVX-
 PT associated disorders, such as cancer, obesity, diabetes and inflammatory
 PT or CNS diseases.
 XX Claim 20; SEQ ID NO 125; 250pp; English.
 PS This invention relates to a novel isolated NOVX polypeptide comprising a
 CC fully defined sequence of, a mature form, one or more conservative
 CC substitutions or at least 95% identity to 247 amino acids as given in the
 CC specification. The invention may be useful for the development of
 CC compounds with a cytostatic, antidiabetic, anorectic, cerebroprotective,
 CC neuroprotective, antiinflammatory, thymometric or cardiac activity. In
 CC addition, the disclosed sequences may prove useful for gene-therapy or
 CC antisense-therapy. The invention may be useful for the diagnosis and
 CC treatment of disorders associated with aberrant expression or activity of
 CC the NOVX polypeptide, such as cancer, diabetes, obesity, and endocrine,
 CC CNS, cardiovascular and inflammatory disorders. They can also be used in
 CC various detection and screening assays, chromosome mapping, tissue typing
 CC and predictive medicine. The present sequence is that of a gene which
 CC encodes a human NOVX protein of the invention.
 XX Sequence 735 BP; 220 A; 138 C; 183 G; 194 T; 0 U; 0 Other;
 SQ

Query Match 43.4%; Score 735; DB 12; Length 735;
 Best Local Similarity 100.0%; Pred. No. 1.2e-188;
 Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 AGTGGCAAGATGGCGTCCCTGGATCGGTCAGGTACTGTGTGGGAGACTCAGGTGTT 69
 DB 1 AGTGGCAAGATGGCGTCCCTGGATCGGTCAGGTACTGTGTGGGAGACTCAGGTGTT 60
 QY 70 GCGAATCTTCCTTAGTCCATCTCTATGCAAAATCAAGTGTGGGAAATCCATCATGG 129
 DB 61 GCGAATCTTCCTTAGTCCATCTCTATGCAAAATCAAGTGTGGGAAATCCATCATGG 120
 QY 130 ACTGTGGGCTGCTCAGTGGATGTGAGATTCATGATTACAAAGAGAACCCAGAGAG 189
 DB 121 ACTGTGGGCTGCTCAGTGGATGTGAGATTCATGATTACAAAGAGAACCCAGAGAG 180
 QY 190 AAGACCTACTACATAGAAATATGGGATGTGGAGGCTCTGTGGCAGTGCAGCAGGTG 249
 DB 181 AAGACCTACTACATAGAAATATGGGATGTGGAGGCTCTGTGGCAGTGCAGCAGGTG 240
 QY 250 AAAAGCACAGACAGCATTTCTCAACTCCGTAATGTTATTTTCGTACAGACTTA 309
 DB 241 AAAAGCACAGACAGCATTTCTCAACTCCGTAATGTTATTTTCGTACAGACTTA 300
 QY 310 ACAATAAGAGAGTCTCCCAAACTTCGCTCGTTGGTCAATGGAGCTCTCAACAGGGAT 369
 DB 301 ACAATAAGAGAGTCTCCCAAACTTCGCTCGTTGGTCAATGGAGCTCTCAACAGGGAT 360
 QY 370 TTGGTGCAACTGGAGTCTTTGGTGACAAATGGGATTTATGATCAAGAACAGTTTGTGAT 429
 DB 361 TTGGTGCAACTGGAGTCTTTGGTGACAAATGGGATTTATGATCAAGAACAGTTTGTGAT 420
 QY 430 AACCAATACCACTGTTGGTAATAGGACCTAACTGGACAGATTCATGAACAAAGGC 489
 DB 421 AACCAATACCACTGTTGGTAATAGGACCTAACTGGACAGATTCATGAACAAAGGC 480
 QY 490 CATGAAGTTTTAACTAGGACTGCTTTCTCGCTGAGGATTTCAATCCAGAGAAATTAAT 549
 DB 481 CATGAAGTTTTAACTAGGACTGCTTTCTCGCTGAGGATTTCAATCCAGAGAAATTAAT 540
 QY 550 TTGAGCTGCAAAATCCACGGTACTTACGTGTCAGGTTCTTCCAAATGCTGTCAAGTCACT 609
 DB 541 TTGAGCTGCAAAATCCACGGTACTTACGTGTCAGGTTCTTCCAAATGCTGTCAAGTCACT 600
 QY 610 AGGTTTTTGTATAGGTCATAGAGAGAGATCTTTTAAAGAGAGAGTATCATCATATCCA 669
 DB 601 AGGTTTTTGTATAGGTCATAGAGAGAGATCTTTTAAAGAGAGAGTATCATCATATCCA 660

QY 670 GGCTTTCTGATCGGAAAGATTGGGGCAGGACATTAAAGAGCCTTCATTATGACTGA 729
DB 661 GGCTTTCTGATCGGAAAGATTGGGGCAGGACATTAAAGAGCCTTCATTATGACTGA 720
QY 730 ATTACACTCATCCTT 744
DB 721 ATTACACTCATCCTT 735
RESULT 6
ADE47765
ID ADE47765 standard; DNA; 739 BP.
AC ADE47765;
XX 29-JAN-2004 (first entry)
DT Human NOV38b gene SEQ ID NO:127.
DE
XX ds; gene; human; cardiant; antiarteriosclerotic; hypotensive;
KW immunosuppressive; dermatological; anorectic; cytostatic; antidiabetic;
KW haemostatic; anti-HIV; antiasthmatic; antibacterial; virucide;
KW neuroprotective; nootropic; antiparkinsonian; antilipaemic; gene therapy;
KW vaccine.
XX
OS Homo sapiens.
XX
PN WO2003076642-A2.
XX
PD 18-SEP-2003.
XX
PF 02-AUG-2002; 2002WO-US024459.
XX
PR 02-AUG-2001; 2001US-0309501P.
PR 03-AUG-2001; 2001US-0310291P.
PR 08-AUG-2001; 2001US-0310951P.
PR 09-AUG-2001; 2001US-0311292P.
PR 13-AUG-2001; 2001US-0311979P.
PR 14-AUG-2001; 2001US-0312203P.
PR 17-AUG-2001; 2001US-0313156P.
PR 17-AUG-2001; 2001US-0313201P.
PR 20-AUG-2001; 2001US-0313702P.
PR 21-AUG-2001; 2001US-0314031P.
PR 23-AUG-2001; 2001US-0314466P.
PR 28-AUG-2001; 2001US-0315403P.
PR 29-AUG-2001; 2001US-0315853P.
PR 31-SEP-2001; 2001US-0316508P.
PR 21-SEP-2001; 2001US-0323936P.
PR 03-DEC-2001; 2001US-0338078P.
PR 05-FEB-2002; 2002US-0354655P.
PR 05-MAR-2002; 2002US-0361764P.
PR 19-APR-2002; 2002US-0373825P.
PR 15-MAY-2002; 2002US-0380971P.
PR 15-MAY-2002; 2002US-0380980P.
PR 16-MAY-2002; 2002US-0381039P.
PR 28-MAY-2002; 2002US-0383761P.
PR 29-MAY-2002; 2002US-0383887P.
PR 01-AUG-2002; 2002US-00210130.
XX (CURA-) CURAGEN CORP.
XX
PI Zerhusen BD, Patturajan M, Kekuda R, Miller CE, Rieger DK;
PI Pena CE, Shinkets RA, Li L, Berghs C, Zhong M, Casman SU, Voss EZ;
PI Boldog FI, Padigaru M, Smithson G, Shenoy SG, Ji W, Gorman L;
PI Vernet CAM, Leite MW, Guo X, Anderson DW, Spytek KA, Gerlach VL;
PI Burgess CE, Khramsov NV, Ort T, Ellerman K, Rastelli L, Agee ML;
PI Chaudhuri A, Chant JS, Dipippo VA, Edinger SR, Eissen A, Gangolli EA;
PI Glot L, Ooi CE, Rothenberg ME, Spaderna SK, Hjalt T, Liu X;
PI Taupier RJ, Catterton E;
XX WPI; 2003-779062/73.
DR P-PSDB; ADE47766.

XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes,
PT atherosclerosis, asthma or AIDS, and in chromosome mapping, tissue typing
PT or pharmacogenomics.
XX
XX Claim 20; SEQ ID NO 127; 562pp; English.
XX
XX The invention relates to a novel (NOVX) human polypeptide. A polypeptide
CC of the invention has cardiant, antiarteriosclerotic, hypotensive,
CC immunosuppressive, dermatological, anorectic, cytostatic, antidiabetic,
CC haemostatic, anti-HIV, antiasthmatic, antibacterial, virucide,
CC neuroprotective, nootropic, antiparkinsonian, and antilipaemic activity.
CC A polynucleotide encoding a polypeptide of the invention may have a use
CC in gene therapy, and as a vaccine. A polypeptide of the invention is
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, the disease selected from a pathology
CC associated with the polypeptide. These may also be used in diagnosing,
CC treating or preventing NOVX-associated disorders such as cardiomyopathy,
CC atherosclerosis, hypertension, scleroderma, obesity, cancer, diabetes,
CC haemophilia, graft-versus-host disease, AIDS, asthma, Crohn's disease,
CC multiple sclerosis, infections, anorexia, cancer-associated cachexia,
CC neurodegenerative disorders (e.g. Alzheimer's disease or Parkinson's
CC disease), haematopoietic disorders, dyslipidaemias and other wasting
CC disorders associated with chronic diseases. The nucleic acids are also
CC used as hybridisation probes, in chromosome mapping, tissue typing,
CC preventive medicine, and pharmacogenomics. The polypeptides are also
CC useful as vaccines. The present sequence encodes a NOVX polypeptide of
CC the invention.
XX
SQ Sequence 739 BP; 222 A; 140 C; 185 G; 192 T; 0 U; 0 Other;
Query Match 43.2%; Score 731.6; DB 10; Length 739;
Best Local Similarity 99.5%; Pred. No. 1e-187;
Matches 734; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 10 AGTGGCAAGATGGCGTCCCTGGATCGGTGAAGTACTGGTGTGGAGACTCAGGTGTT 69
DB 1 AGTGGCAAGATGGCGTCCCTGGATCGGTGAAGTACTGGTGTGGAGACTCAGGTGTT 60
QY 70 GGGAAATCTTCGTTAGTCCATCTCTATGCCAAAATCAAGTGTGGGAAATCCATCATGG 129
DB 61 GGGAAATCTTCGTTAGTCCATCTCTATGCCAAAATCAAGTGTGGGAAATCCATCATGG 120
QY 130 ACTGTGGGCTGCTCAGTGTGATGTCAGAGTTCATGATTACAAAGAGACCCAGAGAG 189
DB 121 ACTGTGGGCTGCTCAGTGTGATGTCAGAGTTCATGATTACAAAGAGAGAACCCAGAGAG 180
QY 190 AAGACCTACTACATAGAAATTTATGGGATGTTGGAGGCTCTGTGGGCAGTGCAGCAGCGTG 249
DB 181 AAGACCTACTACATAGAAATTTATGGGATGTTGGAGGCTCTGTGGGCAGTGCAGCAGCGTG 240
QY 250 AAAAGCACAAGAGCAGTATTTCTACAACTCCGTAATAGTATTTATTTTCGTCACGACTTA 309
DB 241 AAAAGCACAAGAGCAGTATTTCTACAACTCCGTAATAGTATTTATTTTCGTCACGACTTA 300
QY 310 ACAAATAAGAGTCTCTCCCAAAACTTGGCTCGTGTGTCATTGCGAAGCTCTCACAGGGAT 369
DB 301 ACAAATAAGAGTCTCTCTCCCAAAACTTGGCTCGTGTGTCATTGCGAAGCTCTCACAGGGAT 360
QY 370 TTGGTGCCCAACTGGAGTCTTTGGTGACAAATGGGGATTTATGATCAAGAACAGTTTTCGTGAT 429
DB 361 TTGGTGCCCAACTGGAGTCTTTGGTGACAAATGGGGATTTATGATCAAGAACAGTTTTCGTGAT 420
QY 430 AACCAATAACCACTGTTTGGTAAATAGGAGCTTAAACTCGACAGATTCATGAAACAAAGCGC 489
DB 421 AACCAATAACCACTGTTTGGTAAATAGGAGCTTAAACTCGACAGATTCATGAAACAAAGCGC 480
QY 490 CATGAAGTTTAACTAGGACTGCTTTCCTGGCTGAGGATTTCAATCCAGAGAAATTAAT 549
DB 481 CATGAAGTTTAACTAGGACTGCTTTCCTGGCTGAGGATTTCAATCCAGAGAAATTAAT 540
QY 550 TTGGAGTCGACAAATCCAGGTACTTAGCTGCAGGTTTCTTCCAAATGCTGTCAAGCTCAGT 609

```

Db      541  TTGACCTGCAAAATCCAGGTACTTACCTCGAGTTCCTCAATGCTGCAAGCTCAGT   600
Qy      610  AGGTTTTTGTATAAGGTCATAGAGAAGATACATCTTTTAAAGAGAAGGTAAATCAGATTCCA   669
Db      601  AGGTTTTTGTATAAGGTCATAGAGAAGATACATCTTTTAAAGAGAAGGTAAATCAGATTCCA   660
Qy      670  GGCTTTCTGATCGAAAGATTTGGGCGCAGGACATTAAGAGCCTTCATTATGACTGA   729
Db      661  GGCTTTCTGATCGAAAGATTTGGGCGCAGGACATTAAGAGCCTTCATTATGACTGA   720
Qy      730  ATTACACTCATCCTTTGG 747
Db      721  ATTACACTCATCCTAAGG 738

RESULT 7
ID      ADJ79035
XX      ADJ79035 standard; DNA; 739 BP.
AC      ADJ79035;
XX      06-MAY-2004 (first entry)
XX      Human NOVX protein Nov38B gene sequence.
XX      NOVX; cytostatic; antidiabetic; anorectic; cerebroprotective;
XX      neuroprotective; antiinflammatory; thymomimetic; cardiac; gene-therapy;
XX      antisense-therapy; cancer; diabetes; obesity; endocrine disorder;
XX      CNS disorder; cardiovascular disorder; inflammatory disorder;
XX      detection assay; screening assay; chromosome mapping; tissue typing;
XX      predictive medicine; human; Nov38B; gene; ds.
XX      Homo sapiens.
XX      US2004014053-A1.
XX      22-JAN-2004.
XX      01-AUG-2002; 2002US-00210130.
XX      02-AUG-2001; 2001US-0309501P.
XX      03-AUG-2001; 2001US-0310291P.
XX      08-AUG-2001; 2001US-0310951P.
XX      09-AUG-2001; 2001US-0311292P.
XX      13-AUG-2001; 2001US-0311979P.
XX      14-AUG-2001; 2001US-0312203P.
XX      17-AUG-2001; 2001US-0313156P.
XX      17-AUG-2001; 2001US-0313201P.
XX      20-AUG-2001; 2001US-0313643P.
XX      20-AUG-2001; 2001US-0313702P.
XX      21-AUG-2001; 2001US-0314031P.
XX      23-AUG-2001; 2001US-0314466P.
XX      28-AUG-2001; 2001US-0315403P.
XX      29-AUG-2001; 2001US-0315853P.
XX      31-AUG-2001; 2001US-0316508P.
XX      17-SEP-2001; 2001US-0322716P.
XX      21-SEP-2001; 2001US-0323936P.
XX      03-DEC-2001; 2001US-0338078P.
XX      05-FEB-2002; 2002US-0354655P.
XX      05-MAR-2002; 2002US-0361764P.
XX      19-APR-2002; 2002US-0373825P.
XX      15-MAY-2002; 2002US-0380971P.
XX      15-MAY-2002; 2002US-0380980P.
XX      16-MAY-2002; 2002US-0381039P.
XX      28-MAY-2002; 2002US-0383761P.
XX      29-MAY-2002; 2002US-0383887P.
XX      (ZERH/) ZERHUSEN B D.
PA      (PATT/) PATTURAJAN M.
PA      (KEKU/) KEKUDA R.
PA      (MILL/) MILLER C E.
PA      (RIEG/) RIEGER D K.

```

```

PA      (PENA/) PENA C E A.
PA      (SHIM/) SHIMKETS R A.
PA      (LILL/) LI L.
PA      (BERG/) BERGHS C.
PA      (ZHON/) ZHONG M.
PA      (CASM/) CASMAN S J.
PA      (VOSS/) VOSS E Z.
PA      (BOLD/) BOLDOG F L.
PA      (PADI/) PADIGARU M.
PA      (SMIT/) SMITHSON G.
PA      (JIWV/) JI W.
PA      (GORM/) GORMAN L.
PA      (VERN/) VERNET C A M.
PA      (LEIT/) LEITE M W.
PA      (GUOX/) GUO X S.
PA      (ANDE/) ANDERSON D W.
PA      (SPYT/) SPYTEK K A.
PA      (GERL/) GERLACH V.
PA      (BURG/) BURGESS C E.
PA      (KHRA/) KHRAMTSOV N V.
PA      (ORTT/) ORT T.
PA      (ELLE/) ELLERMAN K.
PA      (RAST/) RASTELLI L.
PA      (AGEE/) AGEER M L.
PA      (CHAU/) CHAUDHURI A.
PA      (CHAN/) CHANT J S.
PA      (DIPI/) DIPIPPO V A.
PA      (EDIN/) EDINGER S R.
PA      (EISE/) EISEN A J.
PA      (GANG/) GANGOLLI E A.
PA      (GIOT/) GIOT L.
PA      (OOIC/) OOI C E.
PA      (ROTH/) ROTHENBERG M E.
PA      (SPAD/) SPADERNA S K.
PA      (HJAL/) HJALT T.
PA      (LIUX/) LIU X.
PA      (TAUP/) TAUPIER R J.
PA      (CATT/) CATTERTON E.
PA      (SHEN/) SHENOY S G.
XX      ZERHUSEN BD, Patturajan M, Kekuda R, Miller CE, Rieger DK;
XX      Pena CEA, Shimkets RA, Li L, Berghs C, Zhong M, Casman SJ, Voss EZ;
XX      Boldog FL, Padigaru M, Smithson G, Ji W, Gorman L, Vernet CM;
XX      Leite MW, Guo XS, Anderson DW, Spytek KA, Gerlach V, Burgess CE;
XX      Khrantsov NV, Ort T, Ellerman K, Rastelli L, Agee ML, Chaudhuri A;
XX      Chant JS, Dipippo VA, Edinger SR, Eisen AJ, Gangolli EA, Giot L;
XX      Ooi CE, Rothenberg ME, Spaderna SK, Hjalt T, Liu X, Taupier RJ;
XX      Catterton E, Shenoy SG;
DR      WPI; 2004-108206/11.
DR      P-PSDB; ADJ79036.
XX      New isolated NOVX polypeptides and nucleic acid molecules useful for
XX      treating, preventing and diagnosing pathological conditions with NOVX-
XX      associated disorders, such as cancer, obesity, diabetes and inflammatory
XX      or CNS diseases.
XX      Claim 20; SEQ ID NO 127; 250pp; English.
XX      This invention relates to a novel isolated NOVX polypeptide comprising a
XX      fully defined sequence of, a mature form, one or more conservative
XX      substitutions or at least 95% identity to 247 amino acids as given in the
XX      specification. The invention may be useful for the development of
XX      compounds with a cytostatic, antidiabetic, anorectic, cerebroprotective,
XX      neuroprotective, antiinflammatory, thymomimetic or cardiac activity. In
XX      addition, the disclosed sequences may prove useful for gene-therapy or
XX      antisense-therapy. The invention may be useful for the diagnosis and
XX      treatment of disorders associated with aberrant expression or activity of
XX      the NOVX polypeptide, such as cancer, diabetes, obesity, and endocrine,
XX      CNS, cardiovascular and inflammatory disorders. They can also be used in
XX      various detection and screening assays, chromosome mapping, tissue typing
XX      and predictive medicine. The present sequence is that of a gene which
XX      encodes a human NOVX protein of the invention.

```

XX SQ Sequence 739 BP; 222 A; 140 C; 185 G; 192 T; 0 U; 0 Other;

Query Match 43.2%; Score 731.6; DB 12; Length 739;
Best Local Similarity 99.5%; Pred. No. 1e-187;
Matches 734; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 10 AGTGGCAAGATGCGTCCCTGGATCGGTGCAAGTACTGCTGTTGGAGACTCAGGTGTT 69
DB 1 AGTGGCAAGATCGGTCCTCGATCGGTGCAAGTACTGCTGTTGGAGACTCAGGTGTT 60

QY 70 GGGAAATCTTGGTAGTCCATCTCTATGCCAAATCAAGTCTGGAAATCCATCATGG 129
DB 61 GGGAAATCTTGGTAGTCCATCTCTATGCCAAATCAAGTCTGGAAATCCATCATGG 120

QY 130 ACTGTGGCTGCTCAGTGGATGTCAGAGTTCATGATTACAAAGAAAGAACCCAGAAAG 189
DB 121 ACTGTGGCTGCTCAGTGGATGTCAGAGTTCATGATTACAAAGAAAGAACCCAGAAAG 180

QY 190 AAGACCTACTACATAGAAATTTATGGATGTTGGAGGCTCTGTGGCAGTGCACGCGTG 249
DB 181 AAGACCTACTACATAGAAATTTATGGATGTTGGAGGCTCTGTGGCAGTGCACGCGTG 240

QY 250 AAAAGCACAGAGCAGTATTTCTCAACTCCGTAAATGGTATTATTTTCGTACACGACTTA 309
DB 241 AAAAGCACAGAGCAGTATTTCTCAACTCCGTAAATGGTATTATTTTCGTACACGACTTA 300

QY 310 ACAATAAGAAAGTCTCCCAAACTTCGTCGTCGTTGGTCAATGGAAGTCTCAACAGGGAT 369
DB 301 ACAATAAGAAAGTCTCCCAAACTTCGTCGTCGTTGGTCAATGGAAGTCTCAACAGGGAT 360

QY 370 TTGGTGCACACTGAGTCTTGGTGACAAATGGGGATTTATGATCAAGAACAGTTTGTCTGAT 429
DB 361 TTGGTGCACACTGAGTCTTGGTGACAAATGGGGATTTATGATCAAGAACAGTTTGTCTGAT 420

QY 430 AACCAATACACTGTTGGTAAATAGGACCTAAATCGACAGATTCATGAAACAAAAGCGC 489
DB 421 AACCAATACACTGTTGGTAAATAGGACCTAAATCGACAGATTCATGAAACAAAAGCGC 480

QY 490 CATGAAGTTTAACTAGGACTGCTTCCTCGCTGAGGATTTCAATCCAGAAAGAAATTAAT 549
DB 481 CATGAAGTTTAACTAGGACTGCTTCCTCGCTGAGGATTTCAATCCAGAAAGAAATTAAT 540

QY 550 TTGGACTGCACAAATCCAGGTAAGTCTAGTCCAGGTTCTTCCAACTGTCAGGCTCAGT 609
DB 541 TTGGACTGCACAAATCCAGGTAAGTCTAGTCCAGGTTCTTCCAACTGTCAGGCTCAGT 600

QY 610 AGGTTTTTTGATAGGTCATAGAGAGAGATCTTTTAAAGAGAGGTAATCAGATTCCA 669
DB 601 AGGTTTTTTGATAGGTCATAGAGAGAGATCTTTTAAAGAGAGGTAATCAGATTCCA 660

QY 670 GGCTTCTCTGATCGGAAAGATTTGGGGCAGGAACATTAAGAGCCCTTCAATTATGACTGA 729
DB 661 GGCTTCTCTGATCGGAAAGATTTGGGGCAGGAACATTAAGAGCCCTTCAATTATGACTGA 720

QY 730 ATTACACTCATCTTTGG 747
DB 721 ATTACACTCATCTTAAGG 738

RESULT 8
AAS27286
ID AAS27286 standard; cDNA; 959 BP.

XX AAS27286;
AC
XX
XX
DT 07-NOV-2001 (first entry)
XX cDNA encoding novel signal transduction pathway protein, Seq ID 321.
DE
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
XX anti-inflammatory; anti-HIV; antibacterial; anti-inflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW

organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.
Homo sapiens.
OS
XX
PN WO200154733-A1.
XX
PD
XX
XX
XX
PF 17-JAN-2001; 2001WO-US001312.
XX
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234224P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236357P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 05-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-465460/50.
 DR P-PSDB; AAU17369.
 XX Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders.
 XX Claim 1; SEQ ID NO 321; 880pp; English.
 PS The invention relates to novel isolated polypeptides (I), and
 XX polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative disorders
 CC (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease), chromosomal abnormalities (e.g.
 CC Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.
 CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),
 CC respiratory disorders, dermatological disorders (e.g. wound healing,
 CC epithelial cell proliferation, endocrine disorders (e.g. Addison's
 CC disease), reproductive system disorders, gastrointestinal disorder
 CC (inflammatory disorders), liver disorders (cirrhosis), as stimulators of
 CC B-cell responsiveness to pathogens, activators of T-cells, to induce
 CC higher affinity antibodies, and as a means to induce tumour proliferation
 CC in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-
 CC AAS27850 represent novel signal transduction pathway protein coding
 CC sequences and PCR primers of the invention
 XX

Query Match 37.1%; Score 628.4; DB 4; Length 959;
 Best Local Similarity 96.5%; Pred. No. 1.1e-159;
 Matches 684; Conservative 1; Mismatches 18; Indels 6; Gaps 4;
 QY 8 GCAGTGGCAAGATGGCGTCCCTGGATCGGTGAAGGTACTGGTGTGGGAGACTCAGGTG 67
 DB 1 GCAGTGGCAAGATGGCGTCCCTGGATCGGTGAAGGTACTGGTGTGGGAGACTCAGGTG 60
 QY 68 TTGGGAATCTTGGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAATCCATCAT 127
 DB 61 TTGGGAATCTTGGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAATCCATCAT 120
 QY 128 GGACTGTGGGCTGCTCAGTGGATGTTCAGAGTTCATGATTACAAGAGAGACCCAGAAG 187
 DB 121 GGACTGTGGGCTGCTCAGTGGATGTTCAGAGTTCATGATTACAAGAGAGACCCAGAAG 180
 QY 188 AGAAGACCTACTACATAGAAATTTATGGGATGTGGAGGCTCTGTGGGAGTCCAGCAGCG 247
 DB 181 AGAAGACCTACTACATAGAAATTTATGGGATGTGGAGGCTCTGTGGGAGTCCAGCAGCG 240
 QY 248 TGAAGACCAAGAGCAGTATTCTCAACCTCCGTAATGGTATTATTTTTCGTACACGACT 307
 DB 241 TGAAGACCAAGAGCAGTATTCTCAACCTCCGTAATGGTATTATTTTTCGTACACGACT 300
 QY 308 TAACAATAAGAAAGTCTCCCAAACTTGGGTCTTGGTTCATTTGGAAGCTCTCAACAGGG 367
 DB 301 TAACAATAAGAAAGTCTCCCAAACTTGGGTCTTGGTTCATTTGGAAGCTCTCAACAGGG 360
 QY 368 ATTTGGTCCCAACTGGAGTCTTGGTGCAAAATGGGGATTATGATCAAGAACAGATTTGCTG 427

Db 361 ATTTGGTCCAACTGGAGTCTTGGTGACAAATGGGATATGATCAGAACAGTTTGCTG 420
 QY 428 ATAAACAAATACCACTGTTGGTAAATAGGAGCTAAACTGGACCAAGATTCATGAAACAAAGC 487
 Db 421 ATAAACAAATACCACTGTTGGTAAATAGGAGCTAAACTGGACCAAGATTCATGAAACAAAGC 480
 QY 488 GCCATGAAGTTTAACTAGGACTGCTTCTGGCTGAGGATTTCAATCCAGAAGAAATTA 547
 Db 481 GCCATGAAGTTTAACTAGGACTGCTTCTGGCTGAGGATTTCAATCCAGAAGAAATTA 540
 QY 548 ATTTGACTGTCACAAATCCACGGTACTTAGCTCAGGTTCTTCCAATGCTGTCAAGCTCA 607
 Db 541 ATTTGACTGTCACAAATCCACGGTACTTAGCTCAGGTTCTTCCAATGCTGTCA-CTCA 599
 QY 608 GTAGGTTTTTGTAAAGGTCATAGAGAGATATCTTTTAAAGAGAGTAAATCAGATTC 667
 Db 600 RTAGGTTTTTGTAT-AGGGCATAGAGAGATAC-TTTTAAAGAGAGTAAATCNGATCC 657
 QY 668 CAGGCTTCTGATCGGAAAGATTTGGGCGAGGACATTAAGAGCCT 716
 Db 658 GAGGCTTCTGACTCG---AAAAGATTGGGCGGACCATTAAGAGCCTT 703

RESULT 9
 ADB93464
 ID ADB93464 standard; cDNA; 959 BP.

XX AC ADB93464;
 XX DT 04-DEC-2003 (first entry)
 XX DE Human cDNA encoding a novel protein #311.
 KW ss; gene; human; autoimmune disease; Parkinson's disease; silicosis;
 KW gastrointestinal disease; atherosclerosis; haemophilia; thrombocytopenia;
 KW immunosuppressive agent; adjuvant; enhance immune response;
 KW higher affinity antibody induction;
 KW increased serum immunoglobulin concentration.

XX OS Homo sapiens.

XX PN US2002168711-A1.

XX PD 14-NOV-2002.

XX PF 17-JAN-2001; 2001US-00764868.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 28-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 26-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0220944P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 05-SEP-2000; 2000US-0229513P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.

PA (RUBE/) RUBEN S M.

PA (BARA/) BARASH S C.

PI Rosen CA, Ruben SM, Barash SC;

XX WPI; 2003-719985/68.

DR P-PSDB; ADB94077.

XX New isolated polypeptide useful for diagnosing and treating

PT immunosuppressive conditions such as autoimmune disease and Parkinson's

PT disease.

XX Claim 3; SEQ ID NO 321; 345pp; English.

XX The invention relates to an isolated polypeptide. The polypeptide is

CC useful for diagnosing a pathological condition or a susceptibility to a

CC pathological condition in a subject, by determining the presence or

CC amount of expression of the polypeptide in a biological sample and

CC diagnosing a pathological condition or a susceptibility to a pathological

CC condition based on the presence or amount of expression of the

CC polypeptide. The polypeptide is also useful for identifying a binding

CC partner to the polypeptide, which involves contacting the polypeptide

CC with a binding partner and determining whether the binding partner

CC effects an activity of the polypeptide. The polypeptide or the nucleic

CC acid encoding the polypeptide is useful for preventing, treating, or

CC ameliorating a medical condition, which involves administering the

CC polypeptide or the nucleic acid to a mammalian subject. The nucleic acid

CC is useful for diagnosing a pathological condition or a susceptibility to

CC a pathological condition in a subject, which involves determining the

CC presence or absence of a mutation in the nucleic acid, and diagnosing a

CC pathological condition or susceptibility to a pathological condition

CC based on the presence or absence of the mutation. The polypeptide, the

CC nucleic acid and an antibody to the polypeptide are useful for treating
 CC autoimmune disease, Parkinson's disease, silicosis, gastrointestinal
 CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,
 CC the nucleic acid and the antibody are useful as immunosuppressive agents,
 CC as adjuvants to enhance immune responses, and as agents to induce higher
 CC affinity antibodies and increase serum immunoglobulin concentrations. The
 CC present sequence represents cDNA encoding a novel human protein. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format direct from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20020168711.

XX Sequence 959 BP; 260 A; 200 C; 249 G; 247 T; 0 U; 3 Other;

XX Query Match 37.1%; Score 628.4; DB 10; Length 959;

Best Local Similarity 96.5%; Pred. No. 1.1e-159; Matches 684; Conservative 1; Mismatches 18; Indels 6; Gaps 4;

```
QY      8 GCAGTGGCAAGATGGCGTCCCTGGATCGGGTGAAGGTACTGGTGTGGGAGCTCAGGTG 67
DB      1 GCAGTGGCAAGATGGCGTCCCTGGATCGGGTGAAGGTACTGGTGTGGGAGCTCAGGTG 60
QY     68 TTGGGAAATCTTCCTTATGATCCATCTCTATGCCAAAATCAAGTGTGGGAAATCCATCAT 127
DB     61 TTGGGAAATCTTCCTTATGATCCATCTCTATGCCAAAATCAAGTGTGGGAAATCCATCAT 120
QY     128 GGACTGTGGGCTGCTCAGTGGATGTCAGATTTCAATGATTAACAAGAGGACCCAGAG 187
DB     121 GGACTGTGGGCTGCTCAGTGGATGTCAGATTTCAATGATTAACAAGAGGACCCAGAG 180
QY     188 AGAAGACTTACTACATAGAATTATGGGATGTTGGAGGCTCTGTGGCGAGTGCCAGCAGCG 247
DB     181 AGAAGACTTACTACATAGAATTATGGGATGTTGGAGGCTCTGTGGCGAGTGCCAGCAGCG 240
QY     248 TGAAGAAGCACAAGAGCAGTATCTTACAACTCCGTAATATGTTATTTTGTATCAGCACT 307
DB     241 TGAAGAAGCACAAGAGCAGTATCTTACAACTCCGTAATATGTTATTTTGTATCAGCACT 300
QY     308 TAACAAATAGAAGTCTCTCCAAAACCTTGGTCTGTTGGTCAATGGAGCTCTCAACAGGG 367
DB     301 TAACAAATAGAAGTCTCTCCAAAACCTTGGTCTGTTGGTCAATGGAGCTCTCAACAGGG 360
QY     368 ATTTGGTGCCAACTGGAGTCTTGTGTGACAAATGGGATTTATGATCAAGAACAGTTTGTG 427
DB     361 ATTTGGTGCCAACTGGAGTCTTGTGTGACAAATGGGATTTATGATCAAGAACAGTTTGTG 420
QY     428 ATAAACCAATACCACTGTGTGTATAGGACTAAACTGGACCAATTCATGAACAAAGC 487
DB     421 ATAAACCAATACCACTGTGTGTATAGGACTAAACTGGACCAATTCATGAACAAAGC 480
QY     488 GCCATGAAGTTTAACTAGGACTGCTTCTCGGCTGAGGATTTCAATCCAGAGAAATTA 547
DB     481 GCCATGAAGTTTAACTAGGACTGCTTCTCGGCTGAGGATTTCAATCCAGAGAAATTA 540
QY     548 ATTTGGAGTGCACAAATCCACGGTACTTGTAGCTGAGGTTCTTCCAAATGCTCTCAAGTCA 607
DB     541 ATTTGGAGTGCACAAATCCACGGTACTTGTAGCTGAGGTTCTTCCAAATGCTCTCAAGTCA 599
QY     608 GTAGTTTTTTTGTATAGTCTATAGAGAGAGATATCTTTTAAAGAGAGGTAATCAGATTC 667
DB     600 RTAGTTTTTTTGTAT-AGGGCATAGAGAGAGATAC-TTTTAAAGAGAGGTAATCNGATCC 657
QY     668 CAGGCTTTCTGTATCGGAAAGATTTGGGCGAGGAAACATTTAAAGAGCTT 716
DB     658 GAGGCTTTCTGTACTCG---AAAAGATTGGGCGGAGCAATTTAAAGCGCTT 703
```

RESULT 10
ACH36304
ID ACH36304 standard; cDNA; 504 BP.
XX
AC ACH36304;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human endothelial cell cDNA #4437.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
FN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX

```
PR 30-JUL-2001; 2001US-00918995.  
XX (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX WPI; 2003-615964/58.  
XX  
PT New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
PS Claim 1; SEQ ID NO 23516; 4pp; English.  
XX  
CC The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
SQ Sequence 504 BP; 148 A; 95 C; 132 G; 121 T; 0 U; 8 Other;  
Query Match 26.2%; Score 444; DB 9; Length 504;  
Best Local Similarity 96.4%; Pred. No. 8.6e-110;  
Matches 450; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
QY 20 TGGCGTCCCTCGATCGGTGCAAGTACTGTGTGTTGGAGACTCAGGTGTTGGGAAATCTT 79  
DB 37 TTGGNANNNGGCTCGGNNGAAGGACTGAGCTGTGGGAGACTCAGGTGTTGGGAAATCTT 96  
QY 80 GGTAGTGCATCTCTATGTCGCAAAATCAAGTGTGGGAAATCCATCATGACTGTGGCT 139  
DB 97 CGTTAGTTCATCTCTATGCTTAATCAAGTGTGGGAAATCCATCATGACTGTGGCT 156  
QY 140 GCTCAGTGGATGTCAGAGTTTCATGATTTACAAAGAGGAACCCAGAGAGAGACCTACT 199  
DB 157 GCTCAGTGGATGTCAGAGTTTCATGATTTACAAAGAGGAACCCAGAGAGAGACCTACT 216  
QY 200 ACATAGAAATTTATGGGATGTTGGAGGCTCTGTGGGCGAGTGCCAGAGCGTGAAGACCAA 259  
DB 217 ACATATAATTTATGGATGTTGGAGGCTCTGTGGGCGAGTGCCAGAGCGTGAAGACCAA 276  
QY 260 GAGCAGTATTTCTACAACTCCGTAAATGGTATTTATTTTCGTACAGACTTAACAATAAGA 319  
DB 277 GAGCAGTATTTCTACAACTCCGTAAATGGTATTTATTTTCGTACAGACTTAACAATAAGA 336  
QY 320 AGTCCTCCCAAACTTCGCTGCTTGGTTCATTTGGAAGCTCTCAACAGGAGATTTGGTGCCAA 379  
DB 337 AGTCCTCCCAAACTTCGCTGCTTGGTTCATTTGGAAGCTCTCAACAGGAGATTTGGTGCCAA 396  
QY 380 CTGGAGTCTTGGTGACAAATGGGATTTATGATCAAGAACAGTTTCTGCTGATAACCAATAC 439  
DB 397 CTGGAGTCTTGGTGACAAATGGGATTTATGATCAAGAACAGTTTCTGCTGATAACCAATAC 456  
QY 440 CACTGTTGGTAAATAGGAGCTTAAACTGGACCAAGATTCATGAACAAAG 486
```

```

Db      457 CACTGTTGGTAATAGGAGCTAAACTGACCGAGATTTCATGAACAAG 503
RESULT 11
AA41283
ID      AAX41283 standard; cDNA; 385 BP.
XX
AC      AAX41283;
XX
DT      17-JUN-1999 (first entry)
XX
DE      Human secreted protein 5' EST SEQ ID NO:227.
XX
KW      Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW      forensic; gene therapy; chromosome mapping; signal peptide;
KW      upstream regulatory sequence; cytokine activity; cell proliferation;
KW      differentiation; hematopoiesis regulation; tissue growth regulation;
KW      reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW      thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS      Homo sapiens.
XX
PN      WO9906548-A2.
XX
PD      11-FEB-1999.
XX
PF      31-JUL-1998; 98WO-IB001222.
XX
PR      01-AUG-1997; 97US-00905135.
XX
PA      (GEST ) GENSET.
XX
PI      Dumas Milne Edwards J, Duclert A, Lacroix B;
XX
WPI; 1999-153778/13.
DR
DR      P-PSDB; AAY12450.
XX
XX      New nucleic acids encoding human secreted proteins - obtained from cDNA
XX      libraries prepared from e.g. liver, ovary, brain, prostate, kidney, lung,
XX      umbilical cord, placenta and colon tissue.
XX
PS      Claim 1; Page 509; 824pp; English.
XX
CC      AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
CC      human secreted proteins, and encode the proteins given in AAY12261 to
CC      AAY12514, respectively. The proteins given represent the signal peptide
CC      and an N-terminal fragment of a secreted protein. The nucleic acid
CC      sequences can be used for producing secreted human gene products. They
CC      can also be used to develop products for diagnosis and therapy. The
CC      proteins obtained may have cytokine activity, cell
CC      proliferation/differentiation activity, hematopoiesis regulating
CC      activity, tissue growth regulating activity, reproductive hormone
CC      regulating activity, chemotactic/ chemokinetic activity, haemostatic and
CC      thrombolytic activity, receptor/ ligand activity, anti-inflammatory
CC      activity, tumour inhibition activity or other activities. The products
CC      can be used in forensic, gene therapy and chromosome mapping procedures.
CC      The sequences can also be used for obtaining corresponding promoter
CC      sequences. The nucleic acids encoding the signal peptide can be used for
CC      directing extracellular secretion of a polypeptide or the insertion of a
CC      polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ      Sequence 385 BP; 103 A; 74 C; 109 G; 91 T; 0 U; 8 Other;
Query Match      22.1%; Score 374.8; DB 2; Length 385;
Best Local Similarity 97.4%; Pred. No. 4.5e-91;
Matches 374; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY      10 AGTGGCAAGATGGCGTCCCTGGATCGGGTCAAGTACTGCTGTGGAGACTCAGGTGTT 69
Db      2 AGTGGCAAGATGGCGTCCCTGGATCGGGTCAAGTACTGCTGTGGAGACTCAGGTGTT 61
QY      70 GGGAAATCTTCGTAGTCCATCTCTATGCCAAATCAAGTCTGGAAATCCATCATGG 129

```

```

Db      62 GGGAAATCTTCGTAGTCCATCTCCCTATGCCAAATCAAGTCTGGAAATCCATCATGG 121
QY      130 ACTGTGGGCTGCTCAGTGGATGTCAGAGTTTCATGATTACAAAGGAACCCAGAGAG 189
Db      122 ACTGTGGGCTGCTCAGTGGATGTCAGAGTKCATGATTACAAAGGAACCCAGAGAG 181
QY      190 AAGACCTACTACATAGAAATTATGGGATGTTGGAGGCTCTGTGGGCAGTGCACGAGCGTG 249
Db      182 AAGACCTACTACATAGAAATTATGGGATGTTGGAGGCTCTGTGGGCAGTGCACGAGCGTG 241
QY      250 AAAAGCACAAGAGCAGTATTCTACAACTCCGTTAAATGGTATTATTTTCGTACAGACTTA 309
Db      242 AAAAGCACAAGAGCAGTATTCTACAACTCCGTTAAATGGTATTATTTTCGTACAGACTTA 301
QY      310 ACAAAATAAGAGTCCCTCCCAAACTTGCCTCGTTGGTTCATTGGGAAGCTCTCAACAGGAT 369
Db      302 ACVSATGGGAAGTCTCCCAAAAMTTGCGNCGTTGGTTCATTGGGAAGCTCTCAACAGGAT 361
QY      370 TTGGTGCCCAACTGGAGTCTTTGGTG 393
Db      362 TTGGTGCCCAACTGGAGTCTTTGGTG 385
RESULT 12
AAC17946
ID      AAC17946 standard; cDNA; 321 BP.
XX
AC      AAC17946;
XX
DT      06-OCT-2000 (first entry)
XX
DE      Human secreted protein 5' EST, SEQ ID NO: 22021.
XX
KW      Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW      gene therapy; chromosome mapping; ss.
XX
OS      Homo sapiens.
XX
PN      EP1033401-A2.
XX
PD      06-SEP-2000.
XX
PF      21-FEB-2000; 2000EP-00200610.
XX
PR      26-FEB-1999; 99US-0122487P.
XX
PA      (GEST ) GENSET.
XX
PI      Dumas Milne Edwards J, Duclert A, Giordano J;
XX
WPI; 2000-500381/45.
XX
XX      New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX      obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX      diagnostic, forensic, gene therapy and chromosome mapping procedures.
PS      Claim 1; SEQ ID NO 22021; 71pp + Sequence Listing; English.
XX
CC      The present sequence is one of a large number of 5' ESTs derived from
CC      mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC      identified within the present sequence. The 5' ESTs were prepared from
CC      total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC      sequences usually correspond mainly to the 3' untranslated region (UTR)
CC      of the mRNA because they are often obtained from oligo-dT primed cDNA
CC      libraries. Such ESTs are not well suited for isolating cDNA sequences
CC      derived from the 5' ends of mRNAs and even in those cases where longer
CC      cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC      ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC      to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC      diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC      They are used to obtain upstream regulatory sequences and to design
CC      expression and secretion vectors
XX

```



```
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX P-PSDB; ABB64204.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX Claim 1; SEQ ID NO 19403; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signaling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (AB857737-
XX AB872072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 919 BP; 226 A; 258 C; 257 G; 178 T; 0 U; 0 Other;
XX
XX Query Match 3.9%; Score 66.8; DB 4; Length 919;
XX Best Local Similarity 52.0%; Pred. No. 2.2e-07;
XX Matches 183; Conservative 0; Mismatches 157; Indels 12; Gaps 1;
XX
XX 17 AGATGGCGTCCCTGGATCGGTGAAGGTACTGTGGTGGGAGACTCAGGTGTGGGAAAT 76
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 14 AATGGCGATGATATCGAGTGGGATTTGTGGTGGGATTCGGTGTGGGAGAGA 73
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 77 CTTCGTAGTCCATCTCCTATGCGCAAAATCAAGTGTGGGAAATCCATCATGACATGTGG 136
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 74 CCTCCTGAGCGACCTGATCACCCACCAAGAGCCCTCATCCGCGCCGCGTGGACGTTGG 133
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 137 GCTGCTCAGTGGATGTCAGAGTTCATGATTTACAAAGAGGAACCCAGAGAGAGACCT 196
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 134 GCTGCAACATCCAGTGAAGATGATCCGTTACAGGAGGGACCGTCGGAGTGTCCCT 193
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 197 ACTACATAGAAATATATGGATGTTGGAGGCTCTGTGGGCGAGTGCAGCAGCGTGAAGAAGCA 256
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 194 ACTTCGTGGAGCTTTCGATGTTGGGGATCGCTG-----AACCAAGAGCA 241
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 257 CAAGAGCAGTATTTCAACATCCGTAATATGATATTTTCGTACAGCATTTACAAATA 316
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 242 CGCGCAGCGTCTTATATCGCGGATCATGATGAATCATTTGTGTGACGACCTTACCAACG 301
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 317 AGAAGTCTCCCAAACTTCGTCGTTGGTTCATTTGGAGCTCTCAACAGGA 368
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 302 CCAAGTCGACAGGACGCTAATCGACTGGCTGTACGAGATCGTCAACAGGA 353
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX
XX RESULT 15
XX ABN37775
XX ID ABN37775 standard; DNA; 60 BP.
XX
XX AC ABN37775;
XX
XX DT 15-JUL-2002 (first entry)
```

Human spliced transcript detection oligonucleotide SEQ ID NO:10523.

Human; mouse; rat; splice transcript; detection; RNA transcript;
splice variant; transcriptome; oligonucleotide library; ss.

Homo sapiens.

WO200210449-A2.

07-FEB-2002.

20-JUL-2001; 2001WO-IB001903.

28-JUL-2000; 2000US-0221607P.

02-MAY-2001; 2001US-0287724P.

(COMP-) COMPUGEN INC.

Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
WPI; 2002-257383/30.

New oligonucleotide libraries comprising oligonucleotides which
selectively hybridize to mRNAs transcribed from a transcription unit of a
genome, useful for detecting tissue-, pathology-, and developmental-
specific genes.

Example 1; SEQ ID NO 10523; 47pp; English.

The present invention describes oligonucleotide libraries for detecting
messenger RNAs that populate a (sub-)transcriptome, where the (sub-
)transcriptome comprises messenger RNAs transcribed from multiple
transcription units that populate a genome. The library comprises several
oligonucleotides, each capable of hybridising selectively to a set of
messenger RNAs transcribed from a given transcription unit of the genome,
which encodes one or more messenger RNA splice variants. The
oligonucleotide libraries are useful for detecting mRNAs from a
biological sample, in expression profiling studies, in qualitatively or
quantitatively characterising the corresponding transcriptome, and in
detecting RNA transcripts and splice variants of human or animal
transcriptomes. The libraries may also be used as specialised mini
libraries to detect transcripts of a sub-transcriptome under a particular
biological or pathological state, and so allowing the detection of tissue
- and pathology-specific genes such as those genes only expressed in
specific tissue under a specific pathological condition; to detect
developmental specific genes; and to detect RNA transcripts and splice
variants of a transcriptome of a patient suffering from a particular
disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
rats, humans and mice, which are used in the exemplification of the
present invention. N.B. The sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 60 BP; 7 A; 16 C; 13 G; 24 T; 0 U; 0 Other;

Query Match 3.5%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 891 GGGAGGTCCAAACTTTGTCCCTCTCTGTGTCTCTTACCTTTCTGTCTGTATAG 950
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GGGAGGTCCAAACTTTGTCCCTCTCTGTGTCTCTTACCTTTCTGTCTGTATAG 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Search completed: June 30, 2005, 14:23:50
Job time : 1027.81 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 13:44:48 ; Search time 317.669 Seconds
(without alignments)
8725.600 Million cell updates/sec

Title: US-09-945-173-1

Perfect score: 1694

Sequence: 1 agggagggcagtggaagat.....aaacatgaattattgaacta 1694

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
 - 2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
 - 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
 - 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
 - 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
 - 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	321	18.9	321	4	US-09-513-999C-22021
2	66.8	3.9	1087	4	US-09-270-767-13908
3	51.8	3.1	1141	4	US-09-806-708B-22
4	42.8	2.5	832	4	US-09-621-976-2813
5	42.8	2.5	1141	4	US-09-806-708B-22
6	42	2.5	601	4	US-09-949-016-51128
7	41.8	2.5	832	4	US-09-621-976-2813
8	41.2	2.4	2239	4	US-09-710-279-4411
9	41.2	2.4	2259	4	US-09-710-279-4078
10	41.2	2.4	3475	4	US-09-710-279-4032
11	40.8	2.4	118923	4	US-09-949-016-13227
12	40.6	2.4	264665	4	US-09-949-016-13747
13	39.8	2.3	2273	3	US-09-714-918-40
14	39.8	2.3	2273	3	US-09-265-315-40
15	39.8	2.3	2273	3	US-09-265-315-40
16	39.8	2.3	2273	3	US-09-266-417-40
17	39.8	2.3	2273	4	US-09-528-709-40
18	39.8	2.3	2273	4	US-09-527-745-40
19	39.6	2.3	601	4	US-09-949-016-130492
20	39.6	2.3	3001	4	US-09-539-333D-170
21	39.6	2.3	197875	4	US-09-949-016-15425
22	39.2	2.3	40465	4	US-09-949-016-12561
23	38.4	2.3	300598	4	US-09-949-016-11868
24	38.4	2.3	302604	4	US-09-949-016-14588
25	38.4	2.3	302604	4	US-09-949-016-14589
26	38.4	2.3	308362	4	US-09-949-016-17119
27	38	2.2	601	4	US-09-949-016-206666

ALIGNMENTS

RESULT 1

US-09-513-999C-22021

; Sequence 22021, Application US/09513999C

; Patent No. 6783961

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Duclert, A.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; FILE REFERENCES: 59.US2.REG

; CURRENT APPLICATION NUMBER: US/09/513,999C

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/122,487

; PRIOR FILING DATE: 1999-02-26

; NUMBER OF SEQ ID NOS: 36681

; SOFTWARE: Patent.pm

; SEQ ID NO 22021

; LENGTH: 321

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-513-999C-22021

Query Match 18.9% Score 321; DB 4; Length 321;

Best Local Similarity 100.0%; Pred. No. 4e-86;

Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 871 ATGAAGGAAGTGACATATCGGAGGTCCTTCTCTGTTCTCTGTTCTCTTAC 930

Db 1 ATGAAGGAAGTGACATATCGGAGGTCCTTCTCTGTTCTCTGTTCTCTTAC 60

Qy 931 CTTTCTGCTCTGTATAGATATGTAAGGCTTGTAAATATGATGTTGTCAA 990

Db 61 CTTTCTGCTCTGTATAGATATGTAAGGCTTGTAAATATGATGTTGTCAA 120

Qy 991 ATGATGAGTAATAGCAATGACAGTACTGACAGAGAAATTTACTTCTGCTAGAAC 1050

Db 121 ATGATGAGTAATAGCAATGACAGTACTGACAGAGAAATTTACTTCTGCTAGAAC 180

Qy 1051 TGGAGGGTTTTTATGGGTCTGTATTTTCCACACTCATTTGCTGAAAGCTTAATTAAGTA 1110

Db 181 TGGAGGGTTTTTATGGGTCTGTATTTTCCACACTCATTTGCTGAAAGCTTAATTAAGTA 240

Qy 1111 CTTCAAAAAGTATCTCCATTTTACCTTCTTGGAGGGAACGGTCTTGTATACCGACC 1170

Db 241 CTTCAAAAAGTATCTCCATTTTACCTTCTTGGAGGGAACGGTCTTGTATACCGACC 300

Qy 1171 CTGAGTTGTCTACCCCAACA 1191

Db 301 CTGAGTTGTCTACCCCAACA 321

Sequence 17593, A
Sequence 56497, A
Sequence 206667, A
Sequence 14, Appl
Sequence 12565, A
Sequence 17064, A
Sequence 11808, A
Sequence 13388, A
Sequence 168805, A
Sequence 13258, A
Sequence 12505, A
Sequence 14207, A
Sequence 94845, A
Sequence 94846, A
Sequence 1, Appli
Sequence 1, Appli
Sequence 14498, A

```
RESULT 2
US-09-270-767-13908
; Sequence 13908, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13908
; LENGTH: 1087
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-13908

Query Match      3.9%; Score 66.8; DB 4; Length 1087;
Best Local Similarity 52.0%; Pred. No. 2.2e-09;
Matches 183; Conservative 0; Mismatches 157; Indels 12; Gaps 1;

QY 17 AGATGGCGTCCCTGGATCGGCTGAAGTACTGTTGGGAGACTCAGGTGTTGGGAAT 76
Db 267 AATGGCGATGAATATCGAGTGGGAATGTTGTTGGTGGCGATTGCGGTGGGGAAGA 326

QY 77 CTTGCTAGTCCATCTCTATGCCAAATCAAGTGTCTGGGAATCCATCATGACTGTGG 136
Db 327 CCTCCCTGACGACCTGATCAACCCACACGAAGCCCTCATCCGCGCGGTGGACGTGG 386

QY 137 GCTGCTCAGTGGATGTCAGAGTTCATGATTACAAAGAGAACCCACGAAGAAGACCT 196
Db 387 GCTGCAACATCCAGGTGAAGATGCATCCGTTCCAGGGAGGGCACCGTCCGAGTGTCCCT 446

QY 197 ACTACATAGAAATATGGATGTTGGAGGCTCTGTGGGCGAGTCCAGCAGCGTCAAGAACA 256
Db 447 ACTCTGGAGCTGTTCGATGTTGGGGATCGCTG-----AACCAAGAACA 494

QY 257 CAAGACGAGTATCTCAACCTCCGTAATGTTATTTTCGTACACGACTTAACAAATA 316
Db 495 CGCGCAGCGTCTTCTATGCGGCATCGATGAATCATCTGTTGTCAGCAGCTTACCAAG 554

QY 317 AGAAGTCTCCCAAACTTCGTCGTTGTCATGAGCTTCAACAGGGA 368
Db 555 CCAAGTCGACAGGCGAGCTAATCGACTGCTGTACGAGATCGTCAACAAGGA 606

RESULT 3
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: Consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22
```

```
Query Match      3.1%; Score 51.8; DB 4; Length 1141;
Best Local Similarity 11.1%; Pred. No. 7.2e-05;
Matches 84; Conservative 287; Mismatches 382; Indels 7; Gaps 2;

QY 860 TTAATGGAATAATGAAAGGAAGTGACAATACGAGAGGTCACAACTTTGTCCTGCTCT 919
Db 103 WKDMKTAYBMTWNTKWKGTGWRHYWRWRABDVTVDHVVYTAANNATWTTWCMMDKDKRT 162

QY 920 GTGTTCCCTTACCTTTCTGTCCTGTGTATAGATTATGTAAGACCTTGTGTAATATGAG 979
Db 163 RWWKKNNAATGWDGDDTKYHMMNNNCBVTVMVRYKTDWDWBSBKMYTGMWBWKNWSYD 222

QY 980 ATGTTGCTCAAAAT-GATGCAAGTAATGACCAATGACAGTGTACTGCAGAGAAAATTACT 1038
Db 223 VTYVWVWDDMCKRKVRVRVTRGRNRNYWAWBTAAHRRRYNNNGWVTAHRYRWTNNNN 282

QY 1039 CTTGCTCAGAACTGGAGGCTTTTATGCGGTCTGTAAATTTTCCCACTCATTCCTGCTGAAAG 1098
Db 283 NNAKAMCKEAKYGNWRABVNSTCTTWKSKTKVPTSWANNCRAGDANKHKKWKSAA 342

QY 1099 CTTAATTAAGTACTTCAAAAACGTATCTCCATTGTTTACCTTTCTTGAGGGGACCGTCT 1158
Db 343 MGYYNNNNNNNTYKARHBAWVWHSWKKWHAHAHYSRKKWTBYKRTKTMVNNNN 402

QY 1159 TGTAAACAGCCCTGAGTGTCTACCCCAACAATCTCTGTCATTTCCTCAAGATGCAAAA 1218
Db 403 GTTMMKRMWAWYKMDMDMBGTYNNNNNGRTYGTWKWKWNTYKWKANNCKWRADWH 462

QY 1219 TGGTGTATTATTTGTTCTCCACCATGTCACACAGGAATGCCTTAATAATAGCAACCC 1278
Db 463 KTCETHNTTWMKTYNNNCYKWSMNGK-----SHRBAAVYTWVWWRVYAHANN 516

QY 1279 TTGTCCTCCTCTCTCTCTCTGCAAAATGGCTCAGTACTGGAAAGAGCGGACATAATAGC 1338
Db 517 NNDYMWKATCTKYBYCYSKWNYYAAWYTKSSWNTSYRYRWKTNNSWRWSSTRSMGRA 576

QY 1339 CAGAGTTAAATATAACAAATTAATAATACATAGAGAACGCAATACCAGAAAAAAG 1398
Db 577 NNYARABHYGYKWTWBSHBTBHBRAAGAHYMBMYBAKCHCMKAWYKAKKYAGAG 636

QY 1399 AATTCCTGTAATAATGATGTGAAAAATTGACAGCTCCCTCACTTAAAGTTGCTGTATA 1458
Db 637 GSNNNNNNNNNNNNNNATCARDYYAASRWYMANAKWYKYKBAANNAYYTHANNWG 696

QY 1459 TACAGTCTAGGTTTCTGTTGAAATAGGTAGGTTAAATCTAAGACCTGCAAGAGGC 1518
Db 697 CWNNAATDTRRTWKNNNNNAGTWKNNNNNNAKNAKNAKNAKNAKNAKNAKNAKNAKNA 756

QY 1519 AGTGAGAGACATTTACAGCCTCTCTATTTGTTTAAAGAAAAAGTCAACTCCTGA 1578
Db 757 MRGWHADAAAABTTDKRNNGAYTKYTTNNNNNTYRGVVVTTAARDGWANNNNNNNNNN 816

QY 1579 AATGTCCTTAGCTATAATCAGAAAACTAAGAAATATTATT 1618
Db 817 NGMSDMMVWTVWAVANYGTNNNNNNNNNNNAYAWWTKWYTT 856

RESULT 4
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
```


QY 882 TGACATACGGAGGTCACAACTTGTGCTCCCGTCTCTGCTGCTTCCCTTACCTTCTGTCGCC 941
Db 1620 ACATCATATTCTAAGTGGCAWACGCAATTCGGGTTAAATTTTCAATGCGAGTCTTTATCTCAC 1679
QY 942 TGTGTATAGATTATGTAAGCCCTTGTGTAATATGAGATGTTCTCAAAATGATGACAGTA 1001
Db 1680 ATTATTCATATATGATATATCTTTATTTTGAATTTATATTTGACTTAACCTTGATTAAGTA 1739
QY 1002 AATGAGCAATGACAGTGTACTGCGAGAGAAATTTTACTTCTGCTGCTAGAACTGGAGGGTTTT 1061
Db 1740 TAAACTAACTTTCGTTTACTTCAAGTTTAAATCTTATCGAGTGATATTTTCAGATTCTT 1799
QY 1062 TATGGGCTGCTGAATTTTCCACACTCATT 1090
Db 1800 TATCTTTTATAAAATAGCCCTACAATTT 1828

RESULT 15

US-09-265-315-40
; Sequence 40, Application US/09265315
; Patent No. 6187541

GENERAL INFORMATION:

; APPLICANT: Benton, Bret
; APPLICANT: Lee, Ving J.
; APPLICANT: Malouin, Francois
; APPLICANT: Martin, Patrick K.
; APPLICANT: Schmid, Molly B.
; APPLICANT: Sun, Dongxu

; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS

; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS

; TITLE OF INVENTION: TARGET GENES

; NUMBER OF SEQUENCES: 111

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; STREET: Suite 4700

; CITY: Los Angeles

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; MEDIUM TYPE: storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/265,315

; FILING DATE: March 9, 1999

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/714,918

; FILING DATE: September 13, 1996

; APPLICATION NUMBER: 60/009,102

; FILING DATE: December 22, 1995

; APPLICATION NUMBER: 60/003,798

; FILING DATE: September 15, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/DOCKET NUMBER: 240/247

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 40:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2273 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-09-265-315-40

Query Match 2.3%; Score 39.8; DB 3; Length 2273;
Best Local Similarity 46.5%; Pred. No. 0.44;
Matches 125; Conservative 1; Mismatches 143; Indels 0; Gaps 0;
QY 822 CAAAAATCATCTTAAATGCTTACCCCTTCCAGCCTTACCCCTTTAAATGGAATAATGAAAGGAAG 881
Db 1560 CAATAUCACCTTTTATATTTTATCTTAWACTTGATTCATTTATTTTATTTATTTAGTAAAG 1619
QY 882 TGACAATACGGGAGGTCACAACTTTTGCCTGTTCTCTGTTCTTACCTTACCTTCTGTCGCC 941
Db 1620 ACATCATATTCTAAGTGGCAWACGCAATTCGGGTTAAATTTTCAATGCGAGTCTTTATCTCAC 1679
QY 942 TGTGTATAGATTATGTAAGCCCTTGTGTAATATGAGATGTTCTCAAAATGATGACAGTA 1001
Db 1680 ATTATTCATATATGATATATCTTTATTTTGAATTTATATTTGACTTAACCTTGATTAAGTA 1739
QY 1002 AATGAGCAATGACAGTGTACTGCGAGAGAAATTTTACTTCTGCTGCTAGAACTGGAGGGTTTT 1061
Db 1740 TAAACTAACTTTCGTTTACTTCAAGTTTAAATCTTATCGAGTGATATTTTCAGATTCTT 1799
QY 1062 TATGGGCTGCTGAATTTTCCACACTCATT 1090
Db 1800 TATCTTTTATAAAATAGCCCTACAATTT 1828

Search completed: June 30, 2005, 17:00:57
Job time : 325.669 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 16:53:46 ; Search time 1172.77 Seconds
(without alignments)
9059.822 Million cell updates/sec

Title: US-09-945-173-1
Perfect score: 1694
Sequence: 1 agggagggcagtggaagat.....aaacatgaattattgaacta 1694

Scoring table: IDENTITY NUC
Gap 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1694	100.0	1694	9	US-09-945-173-1
2	1664.6	98.3	2398	19	Sequence 1, Appli
3	1358.8	80.2	1417	13	Sequence 34, Appl
4	735	43.4	735	17	Sequence 467, App
5	731.6	43.2	739	17	Sequence 125, App
6	711	42.0	711	9	Sequence 127, App
7	628.4	37.1	959	9	Sequence 3, Appli
					Sequence 321, App

8	444	26.2	504	10	US-09-918-995-23516	Sequence 23516, A	
9	252.8	14.9	387	9	US-09-983-965-2369	Sequence 2369, Ap	
10	60	3.5	60	10	US-09-908-975-10523	Sequence 10523, A	
11	54.6	3.2	1626	20	US-10-425-115-2369	Sequence 2369, Ap	
12	51.4	3.0	664	19	US-10-767-701-25416	Sequence 25416, A	
13	50.2	3.0	1643	18	US-10-425-114-26886	Sequence 26886, A	
14	50.2	3.0	1643	20	US-10-425-115-75255	Sequence 75255, A	
c	15	46.2	2.7	1584	19	US-10-437-963-25042	Sequence 25042, A
16	45.2	2.7	761	19	US-10-437-963-650	Sequence 650, App	
17	44	2.6	6713	15	US-10-311-455-1054	Sequence 1054, Ap	
18	44	2.6	6713	15	US-10-311-455-80	Sequence 80, Appl	
c	19	43.4	2.6	5641	15	US-10-311-455-1369	Sequence 1369, Ap
c	20	42.8	2.5	3673778	16	US-10-312-841-1	Sequence 1, Appli
c	21	42.4	2.5	444	10	US-09-814-353-17561	Sequence 17561, A
c	22	42	2.5	56153	18	US-10-221-714A-519	Sequence 519, App
c	23	41.2	2.4	13131	18	US-10-240-589C-57	Sequence 57, Appl
c	24	41.2	2.4	56153	18	US-10-221-714A-520	Sequence 520, App
c	25	41	2.4	533	20	US-10-425-115-85901	Sequence 85901, A
c	26	40.8	2.4	457	10	US-09-918-995-36980	Sequence 36980, A
c	27	40.8	2.4	495	19	US-10-021-323-4579	Sequence 4579, Ap
c	28	40.8	2.4	176001	17	US-10-210-556-27	Sequence 27, Appl
c	29	40.8	2.4	186739	17	US-10-210-556-19	Sequence 19, Appl
c	30	40.4	2.4	879	20	US-10-425-115-173241	Sequence 173241,
c	31	40.4	2.4	120239	19	US-10-322-281-271	Sequence 271, App
c	32	40.4	2.4	1790242	20	US-10-719-993-6940	Sequence 6940, Ap
c	33	40.2	2.4	551	20	US-10-425-115-40080	Sequence 40080, A
c	34	40.2	2.4	5461	15	US-10-311-455-972	Sequence 972, App
c	35	40.2	2.4	5461	19	US-10-433-793-8	Sequence 8, Appli
c	36	40.2	2.4	3673778	16	US-10-312-841-2	Sequence 2, Appli
c	37	40	2.4	436	20	US-10-425-115-12033	Sequence 12033, A
c	38	39.8	2.3	577	19	US-10-021-323-162	Sequence 162, App
c	39	39.8	2.3	583	19	US-10-021-323-6081	Sequence 6081, Ap
c	40	39.8	2.3	7498	15	US-10-311-455-230	Sequence 230, App
c	41	39.6	2.3	364	10	US-09-918-995-18370	Sequence 18370, A
c	42	39.6	2.3	608	20	US-10-425-115-80992	Sequence 80992, A
c	43	39.6	2.3	3001	17	US-10-147-603-170	Sequence 170, App
c	44	39.6	2.3	6531	15	US-10-311-455-614	Sequence 614, App
45	39.6	2.3	3673778	16	US-10-312-841-2	Sequence 2, Appli	

ALIGNMENTS

RESULT 1
US-09-945-173-1
; Sequence 1, Application US/09945173
; Patent No. US20020127568A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 47324, A NOVEL HUMAN G-PROTEIN AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 38155-20035.00
; CURRENT APPLICATION NUMBER: US/09/945,173
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,293
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1694
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)...(729)
US-09-945-173-1

Query Match 100.0%; Score 1694; DB 9; Length 1694;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 AGGGAAGGCAGTGGCAAGATGGCGTCCCTGGATCGGTGAAGTACTGGTGTGGGAGAC 60

Db 1 AGGGAAGGCAAGTGGCGTCCCTGATCGGCTGAAGTACTCGTGTGGGAGAC 60
QY 61 TCAGGTGTTGGGAAATCTTCGTTAGTCCATCTCTTATGCAAAATCAAGTGTGGGAAAT 120
Db 61 TCAGGTGTTGGGAAATCTTCGTTAGTCCATCTCTTATGCAAAATCAAGTGTGGGAAAT 120
QY 121 CCATCATGAGTCTGGGCTCTCAGTGGATGTGAGAGTTCATGATTACAAAGAGGAAAC 180
Db 121 CCATCATGAGTCTGGGCTCTCAGTGGATGTGAGAGTTCATGATTACAAAGAGGAAAC 180
QY 181 CCAGAAGAGAGCACTACTACATAGAAATTAATGGGATGTTGGAGGCTCTGTGGCAGTGCC 240
Db 181 CCAGAAGAGAGCACTACTACATAGAAATTAATGGGATGTTGGAGGCTCTGTGGCAGTGCC 240
QY 241 AGCAGGTGAAAAGCAAGAGCAGTATTTCTACAACTCCGTAATAATGGTATTAATTCGTA 300
Db 241 AGCAGGTGAAAAGCAAGAGCAGTATTTCTACAACTCCGTAATAATGGTATTAATTCGTA 300
QY 301 CAGGACTTAACAAATGAAGTCTCCCAAACTTGGTGGTGGTCAATTTGGAAGCTCTC 360
Db 301 CAGGACTTAACAAATGAAGTCTCCCAAACTTGGTGGTGGTCAATTTGGAAGCTCTC 360
QY 361 AACAGGATTTGGTGGCAACTGGAGTCTTGGTGACAAATGGGATTAATGATCAAGAACAG 420
Db 361 AACAGGATTTGGTGGCAACTGGAGTCTTGGTGACAAATGGGATTAATGATCAAGAACAG 420
QY 421 TTTGCTGATACCAATACACATGTTGGTAAATAGGACTTAACTGGACAGATTCATGAA 480
Db 421 TTTGCTGATACCAATACACATGTTGGTAAATAGGACTTAACTGGACAGATTCATGAA 480
QY 481 ACAAGGCCATGAGTTTAACTAGGACTGCTTCTGCTGAGGATTTCAATCCAGAA 540
Db 481 ACAAGGCCATGAGTTTAACTAGGACTGCTTCTGCTGAGGATTTCAATCCAGAA 540
QY 541 GAAATTAATTTGGAGTGCACAAATCCACGCTACTTAGCTGCAGGTTCTTCCAAATGCTGTC 600
Db 541 GAAATTAATTTGGAGTGCACAAATCCACGCTACTTAGCTGCAGGTTCTTCCAAATGCTGTC 600
QY 601 AGCTCAGTATGTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 601 AGCTCAGTATGTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 CAGATCCAGGCTTCTGATCGGAAGATTTGGGCGAGGACATTAAGAGCCTTCAT 720
Db 661 CAGATCCAGGCTTCTGATCGGAAGATTTGGGCGAGGACATTAAGAGCCTTCAT 720
QY 721 TATGACTGAATTAACATCATCTCTTTGGAAGAGTGAGCAAGCAGTGGCAGTTTTTCACAGC 780
Db 721 TATGACTGAATTAACATCATCTCTTTGGAAGAGTGAGCAAGCAGTGGCAGTTTTTCACAGC 780
QY 781 TCATCTGCTGTGTTCAATTAATTAACATCATGAGAGTATGATGATGATGATGATGATGAT 840
Db 781 TCATCTGCTGTGTTCAATTAATTAACATCATGAGAGTATGATGATGATGATGATGATGAT 840
QY 841 TACCTTTACGCTTACCTTTTAAATGGAAGATGGAAGAGTGCACAAATACGGAGGTCCA 900
Db 841 TACCTTTACGCTTACCTTTTAAATGGAAGATGGAAGAGTGCACAAATACGGAGGTCCA 900
QY 901 AACTTTGCTCCTGTTCTGTTGTTCTTACCTTTCTGCTGTTGTTGTTGTTGTTGTTGTT 960
Db 901 AACTTTGCTCCTGTTCTGTTGTTCTTACCTTTCTGCTGTTGTTGTTGTTGTTGTTGTT 960
QY 961 AGCTTTGTGTAATATGAGATGTTGTCAAAATGATGAGTAAATGAGCAATGACAGTGTA 1020
Db 961 AGCTTTGTGTAATATGAGATGTTGTCAAAATGATGAGTAAATGAGCAATGACAGTGTA 1020
QY 1021 CTCAGAGAAATTTACTCTTCTGCTAGAACTGGAGGTTTTTATGGTCTGTGTAATTTCC 1080
Db 1021 CTCAGAGAAATTTACTCTTCTGCTAGAACTGGAGGTTTTTATGGTCTGTGTAATTTCC 1080
QY 1081 CACACTCATTTGCTGAAAGCTTAATTAAGTACTTCAAAACGTAATCTCCATTTGTTTACCT 1140

Db 1081 CACACTCATTTGCTGAAAGCTTAATTAAGTACTTCAAAACGTAATCTCCATTTGTTTACCT 1140
QY 1141 TCTTGAAGGGAAGGCTTCTGTTTAAACAGCCCTGAGTGTCTACCCCAAAACAATCTCTGTC 1200
Db 1141 TCTTGAAGGGAAGGCTTCTGTTTAAACAGCCCTGAGTGTCTACCCCAAAACAATCTCTGTC 1200
QY 1201 ATTTTCAAAAGATGCAAAATGGTGTATTAATTAATGTTCTCCACCAATGTGCACACAGGAAT 1260
Db 1201 ATTTTCAAAAGATGCAAAATGGTGTATTAATTAATGTTCTCCACCAATGTGCACACAGGAAT 1260
QY 1261 GCTTAATAATAGCAACCCCTGTTCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
Db 1261 GCTTAATAATAGCAACCCCTGTTCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
QY 1321 AGAGCGGCACTTAATAGCCAGAGTTAAATATATAATTAATAATTAATAATTAATAATAGAGAAC 1380
Db 1321 AGAGCGGCACTTAATAGCCAGAGTTAAATATATAATTAATAATTAATAATTAATAATAGAGAAC 1380
QY 1381 GCAATACCAAGAAAAAGAAATCTGGTAAATGATGTGAAAAATTCAGACGCTCCCTCACT 1440
Db 1381 GCAATACCAAGAAAAAGAAATCTGGTAAATGATGTGAAAAATTCAGACGCTCCCTCACT 1440
QY 1441 CTTAAGGTTGCTGCTATATACAGTCTAGGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500
Db 1441 CTTAAGGTTGCTGCTATATACAGTCTAGGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500
QY 1501 TAAGACCTGCACAGGCGAGTGAGACATTTTACAGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560
Db 1501 TAAGACCTGCACAGGCGAGTGAGACATTTTACAGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560
QY 1561 GGAAGAGTCAACTCTCTGAAATGTCCTTAGCTATAATTCAGAAAACTAAGAAATATATTCT 1620
Db 1561 GGAAGAGTCAACTCTCTGAAATGTCCTTAGCTATAATTCAGAAAACTAAGAAATATATTCT 1620
QY 1621 GTGTCAACAATGTATTTATGGAGAGTAAATAAGTTCCACAGCAACAAAAACAT 1680
Db 1621 GTGTCAACAATGTATTTATGGAGAGTAAATAAGTTCCACAGCAACAAAAACAT 1680
QY 1681 GAATTTTGAACCTA 1694
Db 1681 GAATTTTGAACCTA 1694

RESULT 2

US-10-479-284-34

; Sequence 34, Application US/10479284

; Publication No. US20040158039A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE CORPORATION ; YUE, Henry;

; APPLICANT: LEE, Ernestine A. ; BECHA, Shanya D. ;

; APPLICANT: BAUGHN, Mariah R. ; YAO, Monique G. ;

; APPLICANT: TANG, Y. Tom ; AU-YOUNG, Janice K. ;

; APPLICANT: LAL, Preeti G. ; WARREN, Bridget A. ;

; APPLICANT: DUGGAN, Brendan M. ; TRAN, Oyen K. ;

; APPLICANT: XU, Yuming ; THANGAVELU, Kavitha ;

; APPLICANT: RICHARDSON, Thomas W. ; BANDMAN, Olga ;

; APPLICANT: JONES, Karen Anne ; YANG, Junming ;

; APPLICANT: EMERLING, Brooke M. ; SWARNAKAR, Anita ;

; APPLICANT: LUO, Wen ; CHAWLA, Narinder K. ;

; APPLICANT: AZIMZAI, Yalda ; KHAN, Farrah A. ;

; APPLICANT: LU, Dying Aina M. ; GRIFFIN, Jennifer A. ;

; APPLICANT: LEE, Soo Yeun ; BURFORD, Neil ;

; APPLICANT: ELLIOTT, Vicki S. ; HONCHELL, Cynthia D. ;

; APPLICANT: HB, Ann ; MASON, Patricia M. ;

; APPLICANT: LI, Joana X. ; HAFALIA, April J.A. ;

; APPLICANT: GURURAJAN, Rajagopal

; TITLE OF INVENTION: SECRETED PROTEINS

; FILE REFERENCE: PF-0998 USN

; CURRENT APPLICATION NUMBER: US/10/479,284

; PRIOR FILING DATE: 2003-11-24

; PRIORITY APPLICATION NUMBER: PCT/US02/16234

; PRIOR FILING DATE: 2002-05-21

; PRIOR APPLICATION NUMBER: US 60/293,728

```

; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/297,019
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/299,297
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/300,537
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/301,936
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/362,439
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/363,649
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 60/366,041
; PRIOR FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PERL Program
; SEQ ID NO 34
; LENGTH: 2398
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7475151
; US-10-479-284-34

```

Query Match 98.3%; Score 1664.6; DB 19; Length 2398;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1667; Conservative 0; Mismatches 4; Indels 0;

17	AGATGGCGTCCCTGGATCGGATGAAGGTACTGTGGTTGGGAGACTCAGGTGTGGGAAT	76
1	AGATGGCGTCCCTGGATCGGATGAAGGTACTGTGGTTGGGAGACTCAGGTGTGGGAAT	60
77	CTTCGTTAGTCCATCTCCTATGSCCAAATCAAGTGTCTGGGAAATCCATCATGCGACTGTGG	136
61	CTTCGTTAGTCCATCTCCTATGSCCAAATCAAGTGTCTGGGAAATCCATCATGCGACTGTGG	120
137	GCTGCTCAGTGGATGTCCAGAGTTTCATGATTACAAAGAAAGGAACCCCGAAGGAAGACCT	196
121	GCTGCTCAGTGGATGTCCAGAGTTTCATGATTACAAAGAAAGGAACCCCGAAGGAAGACCT	180
197	ACTACATAGAAATTATGGATGTTGGAGGCTCTGTGGGCAGTGCAGCAGCGCTGAAAAGCA	256
181	ACTACATAGAAATTATGGATGTTGGAGGCTCTGTGGGCAGTGCAGCAGCGCTGAAAAGCA	240
257	CAAGAGCAGTATTCTCAAACTCCGTAATGCTATTATTTTGGTACACGACTTAAACAAATA	316
241	CAAGAGCAGTATTCTCAAACTCCGTAATGCTATTATTTTGGTACACGACTTAAACAAATA	300
317	AGAACTCCTCCAAAACTTGGTCGTGGTCAATTGGAAAGCTCTCAACAGGGATTGGTGC	376
301	AGAACTCCTCCAAAACTTGGTCGTGGTCAATTGGAAAGCTCTCAACAGGGATTGGTGC	360
377	CAACTGGAGTCTTCGTGACAAATGGGGATTATGATCAAGAAACAGTTGCTGTAATACCCAA	436
361	CAACTGGAGTCTTCGTGACAAATGGGGATTATGATCAAGAAACAGTTGCTGTAATACCCAA	420
437	TACCACCTGTTGGTAATAGGGACTAAACTGGACCAGATTTCATGAAAACAAAGCCCATGAAG	496
421	TACCACCTGTTGGTAATAGGGACTAAACTGGACCAGATTTCATGAAAACAAAGCCCATGAAG	480
497	TTTTTAACTAGAGCTGCTTTCTCGCTGAGGATTTTCAATCCAGAGAAATTAATTTGGACT	556
481	TTTTTAACTAGAGCTGCTTTCTCGCTGAGGATTTTCAATCCAGAGAAATTAATTTGGACT	540
557	GCACAAATCCACGGTACTTTAGCTGCAAGTCTTCCCAATGCTGCTCAAGCTCAGTAGGTTTT	616
541	GCACAAATCCACGGTACTTTAGCTGCAAGTCTTCCCAATGCTGCTCAAGCTCAGTAGGTTTT	600
617	TTGATAAGGTCATAGAGAAGATATCTTTTTAAGAGAAGGTAATCAGATTCCAGGCTTTTC	676
601	TTGATAAGGTCATAGAGAAGATATCTTTTTAAGAGAAGGTAATCAGATTCCAGGCTTTTC	660

Qy	677	CTGATCGGAAAAGATTGGGCGAGAACATTAAGAAGCCCTTCATATAGACTGAATTACAC	736
Db	661	CTGATCGGAAAAGATTGGGCGAGAACATTAAGAAGCCCTTCATATAGACTGAATTACAC	720
Qy	737	TCATCCTTTGGAAAGAGTGAAGCAAGCAGTGGCAGTGTTCACAGCTCATCTTGCTGTGTTTC	796
Db	721	TCATCCTTTGGAAAGAGTGAAGCAAGCAGTGGCAGTGTTCACAGCTCATCTTGCTGTGTTTC	780
Qy	797	AATTAATTACATCAGAGCCCTTTTAAACAAAATCATCTTAAAAATGCTACCCCTTCAGCCCTTAC	856
Db	781	AATTAATTACATCAGAGCCCTTTTAAACAAAATCATCTTAAAAATGCTACCCCTTCAGCCCTTAC	840
Qy	857	CCTTTAATGGAAAAATGAAAGGAAGTGACAAATACGGGAGGTCCAAAATTTTGTCCTCCCTGTTTC	916
Db	841	CCTTTAATGGAAAAATGAAAGGAAGTGACAAATACGGGAGGTCCAAAATTTTGTCCTCCCTGTTTC	900
Qy	917	TCGTGTGTCCTTACCTTTCTGTCCCTGTGTATAGATTATGTAAAAAGCCTTGTGTAAATAT	976
Db	901	TCGTGTGTCCTTACCTTTCTGTCCCTGTGTATAGATTATGTAAAAAGCCTTGTGTAAATAT	960
Qy	977	GAGATGTTGTCAAAATGATGCAGTAATGAGCAATGACAGTGTACTGCAGAGAAAATTTTA	1036
Db	961	GAGATGTTGTCAAAATGATGCAGTAATGAGCAATGACAGTGTACTGCAGAGAAAATTTTA	1020
Qy	1037	CTCTTCCTTAGAACTCGAGGGTTTTTATGGGTCGTGTAAATTTTCCCAACACTCATTTGCTGAA	1096
Db	1021	CTCTTCCTTAGAACTCGAGGGTTTTTATGGGTCGTGTAAATTTTCCCAACACTCATTTGCTGAA	1080
Qy	1097	AGCTTAATTAAGPACTTCAAAAACGATATCTCCATTTGTTTTTACCTCTTGAGGGGAAACGGT	1156
Db	1081	AGCTTAATTAAGPACTTCAAAAACGATATCTCCATTTGTTTTTACCTCTTGAGGGGAAACGGT	1140
Qy	1157	CTTGTTTAAACGAGCCCTGAGTGTCTACCCCAACAACTCTCTGCATTTTCAAGATGCAA	1216
Db	1141	CTTGTTTAAACGAGCCCTGAGTGTCTACCCCAACAACTCTCTGCATTTTCAAGATGCAA	1200
Qy	1217	AATGGTGTTATTATTAATTGTCTCCACCATTTGTACACACAGGAATGCTAATAATACGAAC	1276
Db	1201	AATGGTGTTATTATTAATTGTCTCCACCATTTGTACACACAGGAATGCTAATAATACGAAC	1260
Qy	1277	CCTTGTCTCCCTCTCTCTCTTTGCAAAATGGCTCAGTCACTGGAAAGCGGAGCTAATA	1336
Db	1261	CCTTGTCTCCCTCTCTCTCTTTGCAAAATGGGTCAGTGACTGGAAGAGCGGACCTAATA	1320
Qy	1337	GCAGAGTTTAAATATAATACAAATTAATAATACATAGAGNACAGCAATACCCAGAAAAA	1396
Db	1321	GCAGAGTTTAAATATAATACAAATTAATAATACATAGAGNACAGCAATACCCAGAAAAA	1380
Qy	1397	AGAAATCTGGTAAAAATGATGTGAAAAATTGACAGCTCCCTCACTCTTAAGGTTTGTGCTA	1456
Db	1381	AGAAATCTGGTAAAAATGATGTGAAAAATTGACAGCTCCCTCACTCTTAAGGTTTGTGCTA	1440
Qy	1457	TATACAGTCTAGGTTTTCTGTTTTGGAAAAATAGGTAGGGTAAAAATCTAAGACCTGCACAAGG	1516
Db	1441	TATACAGTCTAGGTTTTCTGTTTTGGAAAAATAGGTAGGGTAAAAATCTAAGACCTGCACAAGG	1500
Qy	1517	GCAGTGCAGACATTTTACAGCCCTCTCTCTATTTGTTTTTTTAAAGGAAAGTCAACTCCT	1576
Db	1501	GCAGTGCAGACATTTTACAGCCCTCTCTCTATTTGTTTTTTTAAAGGAAAGTCAACTCCT	1560
Qy	1577	GAAATGTCCTTAGCTATAATCAGAAAAATTAAGAAATATTATTTCTGTGTCAACAATGTATT	1636
Db	1561	GAAATGTCCTTAGCTATAATCAGAAAAATTAAGAAATATTATTTCTGTGTCAACAATGTATT	1620
Qy	1637	TATGAGAGAGTAAAAATAAGTTTCCACAGCAACACAAAAACATGAATTAT	1687
Db	1621	TATGAGAGAGTAAAAATAAGTTTCCACAGCAACACATTTACATGAATTAT	1671

RESULT 3
US-10-044-090-467
; Sequence 467, Application US/10044090

QY	250	AAAAGCAACAGAGCAGTATTCTCAAACTCCGTAATGGTATTATTTTCGTACACGACTTA	309
DB	241	AAAAGCAACAGAGCAGTATTCTCAAACTCCGTAATGGTATTATTTTCGTACACGACTTA	300
QY	310	ACAAATTAAGAGTCTCCCAAACTTGGCTGGTGGTCAATTTGGAGAGCTCTCAACAGGGAT	369
DB	301	ACAAATTAAGAGTCTCCCAAACTTGGCTGGTGGTCAATTTGGAGAGCTCTCAACAGGGAT	360
QY	370	TTGGTGGCAACTGGAGTCTTGGTGACAAATGGGGATTATGATCAAGAAACAGTTTCTCTGAT	429
DB	361	TTGGTGGCAACTGGAGTCTTGGTGACAAATGGGGATTATGATCAAGAAACAGTTTCTCTGAT	420
QY	430	AACCAATATACACTTGTGGTAAATAGGGAATAAATGGACACAGATTATGAAACAAAGGCG	489
DB	421	AACCAATATACACTTGTGGTAAATAGGGAATAAATGGACACAGATTATGAAACAAAGGCG	480
QY	490	CATCAAGTTTTAACTTAGGACTGCTTCTCTGGCTGAGGATTTCAATCCAGAGAAATTAAT	549
DB	481	CATCAAGTTTTAACTTAGGACTGCTTCTCTGGCTGAGGATTTCAATCCAGAGAAATTAAT	540
QY	550	TTGACTGACAAATCCACGGTACTTAGCTGACAGTCTTCCAAATGGCTGTCAAGCTCAGT	609
DB	541	TTGACTGACAAATCCACGGTACTTAGCTGACAGTCTTCCAAATGGCTGTCAAGCTCAGT	600
QY	610	AGGTTTTTTTGATAAGGTCATAGAGAAGATACTTTTTAAGAGAGAGTAATCAGATTCCA	669
DB	601	AGGTTTTTTTGATAAGGTCATAGAGAAGATACTTTTTAAGAGAGAGTAATCAGATTCCA	660
QY	670	GGCTTCTGTATCGAAAGAATTTGGGGCAGAGACATTAAGAGAGCTTCATATGACTGA	729
DB	661	GGCTTCTGTATCGAAAGAATTTGGGGCAGAGACATTAAGAGAGCTTCATATGACTGA	720
QY	730	ATTACACTCATCCTT 744	
DB	721	ATTACACTCATCCTT 735	

RESULT 5

US-10-210-130-127

; Sequence 127, Application US/10210130

; Publication No. US20040014053A1

; GENERAL INFORMATION:

; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Patturajan, Meera

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Miller, Charles E.

; APPLICANT: Rieger, Daniel K.

; APPLICANT: Pena, Carol E.A.

; APPLICANT: Shimkets, Richard A.

; APPLICANT: Li, Li

; APPLICANT: Berghs, Constance

; APPLICANT: Zhong, Mei

; APPLICANT: Casman, Stacie J.

; APPLICANT: Voss, Edward Z.

; APPLICANT: Boldog, Ferenc L.

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Smithson, Glennda

; APPLICANT: Ji, Weizhen

; APPLICANT: Gorman, Linda

; APPLICANT: Vernet, Corine A.M.

; APPLICANT: Leite, Mario W.

; APPLICANT: Guo, Xiaojia Sasha

; APPLICANT: Anderson, David W.

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Gerlach, Valerie

; APPLICANT: Burgess, Catherine E.

; APPLICANT: Khrantsov, Nikolai V.

; APPLICANT: Ort, Tatiana

; APPLICANT: Ellerman, Karen

; APPLICANT: Rastelli, Luca

; APPLICANT: Agee, Michele L.

; APPLICANT: Chaudhuri, Amitabha

```

; APPLICANT: Chant, John S.
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Hjalte, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elina
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C (Cura-716 SMT)
; CURRENT APPLICATION NUMBER: US/10/210,130
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSegList version 0.1
; SEQ ID NO 127
; LENGTH: 739
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10)..(717)
; US-10-210-130-127

Query Match 43.2%; Score 731.6; DB 17; Length 739;
Best Local Similarity 99.5%; Pred. No. 4.1e-191;
Matches 734; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 10 AGTGGCAAGATGCGTCCCTGGATCGGTGAAGTACTGTGTGGAGACTCAGGTGTT 69
Db 1 AGTGGCAAGATGCGTCCCTGGATCGGTGAAGTACTGTGTGGAGACTCAGGTGTT 60
QY 70 GGGAAATCTTCGTAGTCCATCTCTATGCCAAAATCAAGTGTGGGAAATCCATCATGG 129
Db 61 GGGAAATCTTCGTAGTCCATCTCTATGCCAAAATCAAGTGTGGGAAATCCATCATGG 120
QY 130 ACTGTGGGCTGCTCAGTGGATGTGCAGATTCATGATTACAAAGAGAACCCAGAGAG 189
Db 121 ACTGTGGGCTGCTCAGTGGATGTGCAGATTCATGATTACAAAGAGAACCCAGAGAG 180
QY 190 AAGACCTTACTACATAGAAATTTGGATTTGGAGGCTCTCTGGGCGAGTGCAGCGGTG 249
Db 181 AAGACCTTACTACATAGAAATTTGGATTTGGAGGCTCTCTGGGCGAGTGCAGCGGTG 240
QY 250 AAAAGCAAGAGCAGTATTCTCAACTCCGTAATGGTATTATTTTCGTACGACTTA 309
Db 241 AAAAGCAAGAGCAGTATTCTCAACTCCGTAATGGTATTATTTTCGTACGACTTA 300

; APPLICANT: Chant, John S.
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Hjalte, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elina
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C (Cura-716 SMT)
; CURRENT APPLICATION NUMBER: US/10/210,130
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSegList version 0.1
; SEQ ID NO 127
; LENGTH: 739
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10)..(717)
; US-10-210-130-127

Query Match 42.0%; Score 711; DB 9; Length 711;
Best Local Similarity 100.0%; Pred. No. 2e-185;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 ATGGCGTCCCTGGATCGGTGAAGTACTGTGTGGGAGACTCAGGTGTGGGAAATCT 78
Db 1 ATGGCGTCCCTGGATCGGTGAAGTACTGTGTGGGAGACTCAGGTGTGGGAAATCT 60
QY 79 TCGTTAGTCCATCTCTATGCCAAAATCAAGTGTGGGAAATCCATCATCGATGTGGGC 138
Db 61 TCGTTAGTCCATCTCTATGCCAAAATCAAGTGTGGGAAATCCATCATCGATGTGGGC 120
QY 139 TCGTCACTGATGTGCAGATTCATGATTACAAAGAGAACCCAGAGAGAGACTTAC 198
Db 121 TCGTCACTGATGTGCAGATTCATGATTACAAAGAGAACCCAGAGAGAGACTTAC 180
QY 199 TACATAGAATTATGGGATGTGTGGAGGCTCTGTGGGCGAGTGCAGCGCGTGAAGACCA 258
```

181	Db	TACATAGAAATTATGGGATGTTGGAGGCTCTGTGGGCAGTGCACGAGCGTGAAGACACA	240
259	Qy	AGAGCAGTATTTCTCAACTCCGCTAAATGGTATATTTTTCGTACACGACTTAAACAAATAAG	318
241	Db	AGAGCAGTATTTCTCAACTCCGCTAAATGGTATATTTTTCGTACACGACTTAAACAAATAAG	300
319	Qy	AAGTCCTCCAAAACTTGGCTGCTTGGTTCATTTGGGAAGCTCTCAACAGGGATTTGGTGCCA	378
301	Db	AAGTCCTCCAAAACTTGGCTGCTTGGTTCATTTGGGAAGCTCTCAACAGGGATTTGGTGCCA	360
379	Qy	ACTGGAGTCTTGGTGACAAATGGGGATTTATGATCAAGAAACAGTTTGTCTGATAACCAATA	438
361	Db	ACTGGAGTCTTGGTGACAAATGGGGATTTATGATCAAGAAACAGTTTGTCTGATAACCAATA	420
439	Qy	CCACTGTTGGTAAATAGGGACTAAACTGGAACAGATTTCATGAACAAAGCGCCATGAAGTT	498
421	Db	CCACTGTTGGTAAATAGGGACTAAACTGGAACAGATTTCATGAACAAAGCGCCATGAAGTT	480
499	Qy	TTAACTAGGACTGCTTTTCTGGCTGAGGATTTCAATCCAGAAGAAATTAATTTGGACTGCG	558
481	Db	TTAACTAGGACTGCTTTTCTGGCTGAGGATTTCAATCCAGAAGAAATTAATTTGGACTGCG	540
559	Qy	ACAAATCCACGGTACTTAGCTGACGAGTTCTTCCAATGCTTCAAGCTCAGTAGAGTTTCTTT	618
541	Db	ACAAATCCACGGTACTTAGCTGACGAGTTCTTCCAATGCTTCAAGCTCAGTAGAGTTTCTTT	600
619	Qy	GATAAGGTCATAGAGAAGAGATATTTTTTAAGAGAAGGTAAATCAGATTTCCAGGCTTTCTCT	678
601	Db	GATAAGGTCATAGAGAAGAGATATTTTTTAAGAGAAGGTAAATCAGATTTCCAGGCTTTCTCT	660
679	Qy	GATCGGAAAAGATTTGGGGCAGGAAACATTAAGAGCCCTTCATTATGACTGA	729
661	Db	GATCGGAAAAGATTTGGGGCAGGAAACATTAAGAGCCCTTCATTATGACTGA	711

RESULT 7

```

US-09-764-868-321
; Sequence 321, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: ETZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 321
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (652)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (705)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-321

```

	Query Match	37.1%	Score 628.4	DB 9	Length 959
	Best Local Similarity	96.5%	Pred. No. 1.6e-162		
	Matches 684	Conservative 1	Mismatches 18	Indels 6	Gaps 4
Qy	8	GCATGGCAAGATCGGTCCTCGATCGGTTGAAGGTACTGGTGTGGAGACTCAGGTG	67		
Db	1	GCATGGCAAGATCGGTCCTCGATCGGTTGAAGGTACTGGTGTGGAGACTCAGGTG	60		
Qy	68	TTGGGAAATCTTCTTATGTCATCTCTATGCCAAAATCAAGTCTGGGAAATCCATCAT	127		
Db	61	TTGGGAAATCTTCTTATGTCATCTCTATGCCAAAATCAAGTCTGGGAAATCCATCAT	120		

Qy	128	GGACTGTGGGCTGCTCAGTGTGAGTTCAGAGTTCAAGAGGAACCCAGAG	187
Db	121	GGACTGTGGGCTGCTCAGTGTGAGTTCAGAGTTCAAGAGGAACCCAGAG	180
Qy	188	AGAGAGCTTACTACATAGAAATTATGGGATGTTGGAGGCTCTGTGGCGCAGTGCAGCAGG	247
Db	181	AGAGAGCTTACTACATAGAAATTATGGGATGTTGGAGGCTCTGTGGCGCAGTGCAGCAGG	240
Qy	248	TGAAAAGCACAGAGCAGTATTCTACAACTCCGTAAATGGTATTATTTTCGTACACGACT	307
Db	241	TGAAAAGCACAGAGCAGTATTCTACAACTCCGTAAATGGTATTATTTTCGTACACGACT	300
Qy	308	TAAACAATAAGAACTCTCCCAAACCTTCGTCGTTGGTCATTGGAAGCTCTCAACAGG	367
Db	301	TAAACAATAAGAACTCTCCCAAACCTTCGTCGTTGGTCATTGGAAGCTCTCAACAGG	360
Qy	368	ATTTGGTGCCAACTGGAGTCTTGGTGACAAATGGGGATTATGATCAAGAAACAGTTTGGTG	427
Db	361	ATTTGGTGCCAACTGGAGTCTTGGTGACAAATGGGGATTATGATCAAGAAACAGTTTGGTG	420
Qy	428	ATAACCAAATACCACTGTTGGTAAATAGGCACTTAACTCGAACAGATTATGAAACAAAGC	487
Db	421	ATAACCAAATACCACTGTTGGTAAATAGGCACTTAACTCGAACAGATTATGAAACAAAGC	480
Qy	488	GCCATGAAGTTTAACTATGAGTACTGCTTTCTGGCTGAGGATTTTCAATCCAGAGAAATTA	547
Db	481	GCCATGAAGTTTAACTATGAGTACTGCTTTCTGGCTGAGGATTTTCAATCCAGAGAAATTA	540
Qy	548	ATTTTGGAGTCACAAATCCACGGTACTTAGCTGCAAGTTCTTCCAAATGCTGTCAAGCTCA	607
Db	541	ATTTTGGAGTCACAAATCCACGGTACTTAGCTGCAAGTTCTTCCAAATGCTGTCAA-CTCA	599
Qy	608	GTAGGTTTTTTGATAAGGTCATAGAGAGAGATACTTTTTAAGAGAAGTAAATCAGATTTC	667
Db	600	RTAGGTTTTTTGAT-AGGGCATAGAGAGAGATAC-TTTTAAGAGAAGTAAATCNGATCC	657
Qy	668	CAGGTTTTCTGATCGGAAAAGATTTGGGGCAGGAACATTTAAAGAGCCT	716
Db	658	GAGGCTTCTTGACTCG---AAAAGATTGGGCGGAGCCATTAAAGGCGCTT703	

RESULT 8

```

US-09-918-995-23516
; Sequence 23516, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCES: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23516
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(504)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23516

```

Query Match 26.2%; Score 444; DB 10; Length 504;
Best Local Similarity 96.4%; Pred. No. 9.7e-112;
Matches 450; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy 20 TGGCGTCCCTGGATCGGGTGAAGTACTGTGTTTGGGAGACTCAGGTGCTTGGGAAATCTT 79

Db 37 TTGGNANNNGGCTCGGNGAAGGAGCTGACGTGGGAGACTCAGGTGTTGGGAATCTT 96
QY 80 CGTTAGTCCATCTCTCTATGTCACCAAAATCAAGTGTGGGAATCCATCATGAGACTGTGGCT 139
Db 97 CGTTAGTTCATCTCTCTATGTCACCAAAATCAAGTGTGGGAATCCATCATGAGACTGTGGCT 156
QY 140 GCTCAGTGGAGTGCAGAGTTCATGATTACAAAGAGGAACCCAGAGAGAGACACCTACT 199
Db 157 GCTCAGTGGAGTGCAGAGTTCATGATTACAAAGAGGAACCCAGAGAGAGACCTACT 216
QY 200 ACATAGAAATPATGGAGTGTGGAGGCTCTGTGGCAGTGCACAGCGGTGAAAAGCACAA 259
Db 217 ACATATAATPATGGAGTGTGGAGGCTCTGTGGCAGTGCACAGCGGTGAAAAGCACAA 276
QY 260 GACGAGTATCTCAACTCCGTAATGGTATTATTTTCGTACAGACTTAAACAATAAGA 319
Db 277 GAGCAGTATTCTCAACTCCGTAATGGTATTATTTTCGTACAGACTTAAACAATAAGA 336
QY 320 AGTCCTCCCAAACTTGCCTGCTGTTGGTCAATTCGAAGCTCTCAACAGGATTTGGTGCCAA 379
Db 337 AGTCCTCCCAAACTTGCCTGCTGTTGGTCAATTCGAAGCTCTCAACAGGATTTGGTGCCAA 396
QY 380 CTGGAGTCTTTGGTACAAATGGGATTTATGATCAAGAACAGTTTGTCTGATAACCAATAC 439
Db 397 CTGGAGTCTTTGGTACAAATGGGATTTATGATCAAGAACAGTTTGTCTGATAACCAATAC 456
QY 440 CACTGTGGTAATAGGACTAACTGGACCAAGATTCATGAACAAAG 486
Db 457 CACTGTGGTAATAGGACTAACTGGACCAAGATTCATGAACAAAG 503

RESULT 9
US-09-983-965-2369
; Sequence 2369, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 2369
; LENGTH: 387
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 64-LIB3057-021-Q1-K1-H8
US-09-983-965-2369

Query Match 14.9%; Score 252.8; DB 9; Length 387;
Best Local Similarity 84.5%; Pred. No. 5.6e-59;
Matches 333; Conservative 0; Mismatches 52; Indels 9; Gaps 4;
QY 760 GCAGTGGCAGTTTTCACAGCTCATCTTGTGTGT-TCAATTATTACCATCAGCGCTTT 818
Db 1 GCGGTGGCAGTTTTCACAGCTTCTTCTTCTGCTGTGCGATTATTATGTCACGCTTT 60
QY 819 TAACAAATCATC-TTAAATGTACCCCTTACCGCTTACCCCTTTAATGGAAAAATGAAG 877
Db 61 TAATGAATCATCTTTAAATGCCACCTTTACGCTCACCTTTAGTGGAAAACTGAAG 120
QY 878 GAAGTGAACAATACGGGAGGTCCAAACTTTGTCCCTGTCTCTGTGTTCTTACCTTTCTG 937
Db 121 GAAGTGAACAATACGGGAGGTCCAGACTTGTGTC-TCATTCTATTCTTTCACCTTTCTG 176

QY 938 TCCCTGTGTATAGATTATGTAAAGCCTTGTGTAAATATGAGATGTTGTCAAATGATGC 997
Db 177 TCCCTGTGTATAGATTATGTAAAGCCTTGTGTAAATATGAGATGTTGTCAAATGATGC 233
QY 998 AGTAAATGAGCAATGACAGTGTACTGCAGAGAAAAATTTACTCTTGCCTAGAACTGAGGG 1057
Db 234 AGTAAATGAGCAGTGCAGAGATGCTGCAGAGAAAAATTTACTCTTGCCTAGAACTGAGGA 293
QY 1058 TTTTATGGTCTGTAAATTTTCCACACATCATCTGCTGAAAGCTTAATTAAGTACTTCAA 1117
Db 294 TTTTATGGTCTGTAAATTTTCCACACATCATCTGCTGAAAGCTTAATTAAGTACTTCAA 353
QY 1118 AACGTATCTCCATTGTTTACCTCTTGTAGGGGA 1151
Db 354 AATGATCTCCATTGTTTACCTCTTGTAGGGGA 387

RESULT 10
US-09-908-975-10523
; Sequence 10523, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10523
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-10523

Query Match 3.5%; Score 60; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 891 GGGAGGTCCAAACTTTTGTCCCTGTTCTCTGTGTCTTACCTTCTGTCCTGCTGTATAG 950
Db 1 GGGAGGTCCAAACTTTTGTCCCTGTTCTCTGTGTCTTACCTTCTGTCCTGCTGTATAG 60

RESULT 11
US-10-425-115-2369
; Sequence 2369, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 2369
; LENGTH: 1626
; TYPE: DNA
; ORGANISM: Zea mays

```
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_102161C.1
US-10-425-115-2369

Query Match      3.2%; Score 54.6; DB 20; Length 1626;
Best Local Similarity 60.4%; Pred. No. 0.00074;
Matches 90; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 21 GCGTCCTCGATCGGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTC 80
Db 342 GCGCCCTGCGCGCAGGTTCTGTACTCGTGTGGGAGCTCAGGTGTGGGAAATCTTC 401

Qy 81 GTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGAGCTGGGCTG 140
Db 402 ATTGGTGCATCTCTTACTGAAAGATTCTGCACTGGCTGACCGCCCAACAATTGGGTG 461

Qy 141 CTCAGTGCATGTCAGAGTTTCATGATTACA 169
Db 462 TGCAGTTGGTGTAAACATATTACTTACA 490

RESULT 12
US-10-767-701-25416
; Sequence 25416, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5335)B
; CURRENT APPLICATION NUMBER: US/10767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 25416
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30951803
US-10-767-701-25416

Query Match      3.0%; Score 51.4; DB 19; Length 664;
Best Local Similarity 59.1%; Pred. No. 0.0034;
Matches 88; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 21 GCGTCCTCGATCGGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTC 80
Db 398 GCGCCCTGCGCGCAGGTCGTGTCTGTCTGTGGGAGCTCAGGGGTGTGTAATCTTC 457

Qy 81 GTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGAGCTGGGCTG 140
Db 458 ATTGGTGCATCTCTATTGAAAGTTCTGCACTGGCTGACCGCCCAACAATTGGGTG 517

Qy 141 CTCAGTGCATGTCAGAGTTTCATGATTACA 169
Db 518 TGCAGTTGGTGTAAACATATCACTTACA 546

RESULT 13
US-10-425-114-26886
; Sequence 26886, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5335)B
; CURRENT APPLICATION NUMBER: US/10425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26886
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4666-008-F5_FLI
US-10-425-114-26886

Query Match      3.0%; Score 50.2; DB 18; Length 1643;
Best Local Similarity 59.4%; Pred. No. 0.012;
Matches 85; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 26 CCCTGGATCGGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTTAG 85
Db 325 CCTTGGCCCAAGTGGCGTCTCATCTCGGTGATTGAGGTGTGGGAAATCTTCATTGG 384

Qy 86 TCATCTCTATGTCGCAAAATCAAGTGTGGGAAATCCATCATGAGCTGTGGGCTGCTCAG 145
Db 385 TGCATCTCATTTTGAAAAGTTCTGCCATTGTCTGACCATCTCAAACAGTAGGATGCACTG 444

Qy 146 TGGATGTCAGAGTTTCATGATTAC 168
Db 445 TGGGCATTAAACATGTTACTTTAC 467

RESULT 14
US-10-425-115-75255
; Sequence 75255, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 75255
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_168664C.1
US-10-425-115-75255

Query Match      3.0%; Score 50.2; DB 20; Length 1643;
Best Local Similarity 59.4%; Pred. No. 0.012;
Matches 85; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 26 CCCTGGATCGGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTTAG 85
Db 325 CCTTGGCCCAAGTGGCGTCTCATCTCGGTGATTGAGGTGTGGGAAATCTTCATTGG 384

Qy 86 TCATCTCTATGTCGCAAAATCAAGTGTGGGAAATCCATCATGAGCTGTGGGCTGCTCAG 145
Db 385 TGCATCTCATTTTGAAAAGTTCTGCCATTGTCTGACCATCTCAAACAGTAGGATGCACTG 444

Qy 146 TGGATGTCAGAGTTTCATGATTAC 168
Db 445 TGGGCATTAAACATGTTACTTTAC 467

RESULT 15
US-10-437-963-25042/c
; Sequence 25042, Application US/10437963
; Publication No. US20040123343A1
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_102161C.1
US-10-425-115-2369

Query Match      3.2%; Score 54.6; DB 20; Length 1626;
Best Local Similarity 60.4%; Pred. No. 0.00074;
Matches 90; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 21 GCGTCCTCGATCGGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTC 80
Db 342 GCGCCCTGCGCGCAGGTTCTGTACTCGTGTGGGAGCTCAGGTGTGGGAAATCTTC 401

Qy 81 GTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGAGCTGGGCTG 140
Db 402 ATTGGTGCATCTCTTACTGAAAGATTCTGCACTGGCTGACCGCCCAACAATTGGGTG 461

Qy 141 CTCAGTGCATGTCAGAGTTTCATGATTACA 169
Db 462 TGCAGTTGGTGTAAACATATTACTTACA 490

RESULT 12
US-10-767-701-25416
; Sequence 25416, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5335)B
; CURRENT APPLICATION NUMBER: US/10767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 25416
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30951803
US-10-767-701-25416

Query Match      3.0%; Score 51.4; DB 19; Length 664;
Best Local Similarity 59.1%; Pred. No. 0.0034;
Matches 88; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 21 GCGTCCTCGATCGGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTC 80
Db 398 GCGCCCTGCGCGCAGGTCGTGTCTGTCTGTGGGAGCTCAGGGGTGTGTAATCTTC 457

Qy 81 GTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGAGCTGGGCTG 140
Db 458 ATTGGTGCATCTCTATTGAAAGTTCTGCACTGGCTGACCGCCCAACAATTGGGTG 517

Qy 141 CTCAGTGCATGTCAGAGTTTCATGATTACA 169
Db 518 TGCAGTTGGTGTAAACATATCACTTACA 546

RESULT 13
US-10-425-114-26886
; Sequence 26886, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26886
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4666-008-F5_FLI
US-10-425-114-26886

Query Match      3.0%; Score 50.2; DB 18; Length 1643;
Best Local Similarity 59.4%; Pred. No. 0.012;
Matches 85; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 26 CCCTGGATCGGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTTAG 85
Db 325 CCTTGGCCCAAGTGGCGTCTCATCTCGGTGATTGAGGTGTGGGAAATCTTCATTGG 384

Qy 86 TCATCTCTATGTCGCAAAATCAAGTGTGGGAAATCCATCATGAGCTGTGGGCTGCTCAG 145
Db 385 TGCATCTCATTTTGAAAAGTTCTGCCATTGTCTGACCATCTCAAACAGTAGGATGCACTG 444

Qy 146 TGGATGTCAGAGTTTCATGATTAC 168
Db 445 TGGGCATTAAACATGTTACTTTAC 467

RESULT 14
US-10-425-115-75255
; Sequence 75255, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 75255
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_168664C.1
US-10-425-115-75255

Query Match      3.0%; Score 50.2; DB 20; Length 1643;
Best Local Similarity 59.4%; Pred. No. 0.012;
Matches 85; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 26 CCCTGGATCGGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTTAG 85
Db 325 CCTTGGCCCAAGTGGCGTCTCATCTCGGTGATTGAGGTGTGGGAAATCTTCATTGG 384

Qy 86 TCATCTCTATGTCGCAAAATCAAGTGTGGGAAATCCATCATGAGCTGTGGGCTGCTCAG 145
Db 385 TGCATCTCATTTTGAAAAGTTCTGCCATTGTCTGACCATCTCAAACAGTAGGATGCACTG 444

Qy 146 TGGATGTCAGAGTTTCATGATTAC 168
Db 445 TGGGCATTAAACATGTTACTTTAC 467

RESULT 15
US-10-437-963-25042/c
; Sequence 25042, Application US/10437963
; Publication No. US20040123343A1
```

```

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 25042
; LENGTH: 1584
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29969C.1
US-10-437-963-25042

```

```

Query Match      2.7%  Score 46.2; DB 19; Length 1584;
Best Local Similarity 54.4%; Pred. No. 0.15;
Matches 93; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Qy 2 GGAAGGCGAGTGGCAAGATGGCGTCCCTGGATCGGGTGAAGGTACTGGTGTTCGGAGACT 61
Db 1255 GGGACCTCAACGGCGGGCGGCACCCCTGCGGCCAGGTGCGGTGCTCGTGGGGACT 1196

Qy 62 CAGGTGTGGGAATCTTCGTAGTCCATCTCTATGCCAAATCAAGTCTGGGAATC 121
Db 1195 CAGGTGTGGGGAAATCATCATTTGGTTCATCTCATTTGAAAGGCTCTGCAATTGCTCGAC 1136

Qy 122 CATCATGGACTGTGGGCTGCTCAGTGGATGTCAGAGTTTCATGATTACAAAG 172
Db 1135 CACCCCAACGATGGATGGATGGCGAGTGAUGTTAAATATATTTTATTGGAAG 1085

```

Search completed: June 30, 2005, 23:57:27
Job time : 1188.77 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 20:25:54 ; Search time 24 Seconds
(without alignments)
946.132 Million cell updates/sec

Title: US-09-945-173-2
Perfect score: 1231
Sequence: 1 MASLDRKVLVLGDSGVGKS.....PGPPDRKRFGAGTLKSLHYD 236
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	259	21.0	416	2 T31486	hypothetical prote
2	171.5	13.9	217	2 S36365	GTP-binding protei
3	169	13.7	203	2 S51495	GTP-binding protei
4	167	13.6	203	2 B38202	GTP-binding protei
5	166.5	13.5	202	2 J41430	GTP-binding protei
6	164.5	13.4	201	2 J42488	GTP-binding protei
7	164.5	13.4	206	2 J4107	membrane vesicle t
8	163.5	13.3	196	2 PS0279	GTP-binding protei
9	163.5	13.3	258	2 B86153	ARA-5 [imported]
10	162.5	13.2	208	2 A34716	GTP-binding protei
11	162	13.2	221	2 H71444	GTP-binding protei
12	162	13.2	234	2 S38083	probable purine nu
13	161	13.1	203	2 B34716	GTP-binding protei
14	160.5	13.0	203	2 T50323	ypt1-related prote
15	160.5	13.0	206	2 S04590	GTP-binding protei
16	160.5	13.0	209	2 T31551	hypothetical prote
17	160	13.0	205	2 S36368	GTP-binding protei
18	159	12.9	205	2 T03629	GTP-binding protei
19	159	12.9	206	2 S33531	GTP-binding protei
20	158.5	12.9	203	2 S34253	GTP-binding protei
21	158.5	12.9	208	2 A44334	GTP-binding protei
22	158	12.8	201	2 J53337	GTP-binding protei
23	158	12.8	202	2 S38740	GTP-binding protei
24	157.5	12.8	201	2 S06147	GTP-binding protei
25	157.5	12.8	201	2 S39565	GTP-binding protei
26	157	12.8	206	2 T78851	GTP-binding protei
27	157	12.8	206	2 C96529	probable RAB7 GTP
28	156	12.7	201	2 D38625	GTP-binding protei
29	156	12.7	206	2 H96562	hypothetical prote

30	155.5	12.6	200	2 S12790	GTP-binding protei
31	155.5	12.6	208	2 S40207	GTP-binding protei
32	155.5	12.6	208	2 A38202	GTP-binding protei
33	155.5	12.6	218	2 T31656	Ras-related protei
34	155	12.6	205	2 T40425	GTP-binding protei
35	155	12.6	205	2 S38339	GTP-binding protei
36	155	12.6	211	2 S34729	GTP-binding protei
37	155	12.6	216	2 JC4108	GTP-binding protei
38	154.5	12.6	205	1 TVDGYF	GTP-binding protei
39	154.5	12.6	205	1 TVRUYF	GTP-binding protei
40	154.5	12.6	205	1 TVMSYP	GTP-binding protei
41	154.5	12.6	217	2 S30273	GTP-binding protei
42	154	12.5	211	2 T29035	hypothetical prote
43	154	12.5	215	1 TVBYQ4	GTP-binding protei
44	154	12.5	215	2 T14565	GTP-binding protei
45	154	12.5	215	2 S57462	GTP-binding protei

ALIGNMENTS

RESULT 1

T31486
hypothetical protein Y116A8C.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31486
R;McMurray, A.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z21041
A;Accession: T31486
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-416 <WIL>
A;Cross-references: UNIPROT:Q9U2V7; EMBL:AL117204; PIDN:CAB55120.1; CESP:Y116A8C.10
A;Experimental source: clone Y116A8C
C;Genetics:
A;Gene: CESP:Y116A8C.10
A;Introns: 25/2; 63/3; 103/2; 125/1

Query Match 21.0%; Score 259; DB 2; Length 416;
Best Local Similarity 32.2%; Pred. No. 3.8e-14;
Matches 76; Conservative 32; Mismatches 86; Indels 42; Gaps 10;

Qy	1	MASLDR--VKVLVLGDSGVGKSLVHLLCONQVLGNP-----SWTVGCSVDVRVHDYK 51
Db	162	MSSLDSSSTKILVLGDSVCGKTSLCHCIAGG--GEPPGGGRSFDSTIGATVVMWMEYR 218
Qy	52	EGTPEKTYTYIELWDVGGSGVSSVSKSTRVFNYSVNGIIFVHDLTNKSSQNLRRW-- 109
Db	219	AGTPEQRTLELWDIGGMVAHRQAAQ----VFFEGAVGAILVHDLTNKSEENLATWLT 274
Qy	110	SLRALNRDLVPTGLVLTNGDYDQFPADNOIPIIVIGTKLDQI--HS---TKRHEVLTRT 164
Db	275	MLDGKPRGAAPKSKSDPAVALKVDIESCNIPVLIVGTKADLVPHKGPVSVDYDGLGNA 334
Qy	165	AFLAEDFNPEINLDCCTNPRYLAAGSSNAVKLRFDFDKVIEK-----RYFLREG 213
Db	335	-----NSAISVWKPENRRKNSKKPMKIGR--KLIVVEKEKVKFGQYLARG 378

RESULT 2

S36365
GTP-binding protein yptV2 - Volvox carteri
C;Species: Volvox carteri
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: S36365
R;Fabry, S.; Jacobsen, A.; Huber, H.; Palme, K.; Schmitt, R.
Curr. Genet. 24, 229-240, 1993
A;Title: Structure, expression, and phylogenetic relationships of a family of ypt genes
A;Reference number: S36365; MUID:94037148; PMID:8221932
A;Accession: S36365
A;Molecule type: DNA

[illegible]

[illegible]

E;15-131/Domain: translation elongation factor Tu homology <ETU>
F;25-29/Region: nucleotide-binding motif A (P-loop)
F;128-131/Region: GTP-binding NKXD motif
F;158-160/Region: GTP-binding SAK/L motif
F;202,203/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 13.1%; Score 163; DB 2; Length 203;
Best Local Similarity 25.2%; Pred. No. 2.8e-06;
Matches 55; Conservative 39; Mismatches 76; Indels 48; Gaps 9;

Qy 7 VKVLGLDGVGKSLVHLLCONVLGNPSWTVGCSVDVRVHDYKEGTPEKTYVIELWD 66
Db ||:||||| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||
Qy 16 VKLLIGDSGVGKSL--LLRSDGSFTSFATIGIDPKRTIE--LEGKRILQIWD 70
Db ||:||||| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||
Qy 67 VGGSVGSASSVKSTRAVFNSVNGIIFVHDLTNKKSQNLRRMSLEALNRDLVPTGLVT 126
Db ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||
Qy 71 TAGQ---ERFTITTAAYRGAMGILLVYDVTDEKSGSIRNMI---RNI----- 113
Db ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||
Qy 127 NGDYDQEQADNQIPLLVITGTKLDQIHKYKHEVLRTAFIAEDFPENPEINLCTNPRYL 186
Db ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||
Qy 114 -----EQHASVSNKMLIGNCKDMTE--KKVYDSSRGKSLADEY-----GIRFL 155
Db ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||
Qy 187 AAGSSNAVKLSRFF---DKVIEKRYFLREGNQIPGRPD 221
Db ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||
Qy 156 ETSAKNSVNERAFIGLAKDIKKRMI-----DTPNDPD 188
Db ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||

RESULT 14
T50323
yptl-related protein 1 [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50323
R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Cadieu, E.; Lelaure, V.; Galibert, P.
submitted to the ENBL Data Library, January 2000
A:Reference number: Z25061
A:Accession: T50323

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 19:57:06 ; Search time 84 Seconds
(without alignments)
1438.699 Million cell updates/sec

Title: US-09-945-173-2
Perfect score: 1231
Sequence: 1 MASLDKVLVLGDSGVGKS.....PGFPDRKRFAGTLKSLHYD 236

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.4

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1216	98.8	236	2 Q8WUD3	Q8wud3 homo sapien
2	1133	92.0	236	2 Q9D4V7	Q9d4v7 m mus muscu
3	1132	92.0	236	2 Q8BMU2	Q8bm2 m mus muscu
4	1085.5	88.2	235	2 Q6GPM2	Q6gpm2 xenopus lae
5	1026.5	83.4	233	2 Q6TNS7	Q6tns7 brachydanio
6	641	52.1	129	2 Q9D0M6	Q9d0m6 m mus muscu
7	593.5	48.2	236	2 Q7Q5M9	Q7q5m9 anopheles g
8	502	40.8	274	2 Q9VXA9	Q9vxa9 drosophila
9	276	22.4	354	2 Q948D2	Q948d2 oryza sativ
10	276	22.4	354	2 Q7X985	Q7x985 oryza sativ
11	265	21.5	342	1 Y813 ARATH	Q8c5j9 arabidopsis
12	259	21.0	416	2 Q9U2V7	Q9u2v7 caenorhabdi
13	224	18.2	431	2 Q9VIB5	Q9vib5 arabidopsis
14	220.5	17.9	242	2 Q9VQ66	Q9vq66 drosophila
15	189.5	15.4	208	2 Q75CA9	Q75ca9 ashbysa goss
16	178.5	14.5	197	2 Q18337	Q18337 drosophila
17	175.5	14.3	206	2 Q9NFG0	Q9nfg0 plasmodium
18	173.5	14.1	197	2 Q9W4A0	Q9w4a0 drosophila
19	173	14.1	205	2 Q8TGD9	Q8tgd9 emericeila
20	171.5	13.9	217	1 YPT2 VOLCA	Q36861 volvox cart
21	171	13.9	203	2 Q41338	Q41338 lycopersico
22	170.5	13.9	401	2 Q9GNV1	Q9gnv1 leishmania
23	170	13.8	203	2 Q24112	Q24112 nicotiana p
24	169.5	13.8	201	2 Q9W5X0	Q9w5x0 drosophila
25	169	13.7	202	2 Q40203	Q40203 lotus japon
26	169	13.7	203	1 RYL1 YARLI	P41924 yarrowia li
27	168.5	13.7	201	2 Q6P727	Q6p727 xenopus tro
28	168.5	13.7	227	2 Q8AV12	Q8av12 xenopus lae
29	168.5	13.7	231	2 Q6NTJ2	Q6ntj2 xenopus lae
30	167.5	13.6	219	2 Q7Z160	Q7z160 toxoplasma
31	167	13.6	203	1 YPT2_MAZE	Q05737 zea mays (m

Query Match

98.8%; Score 1216; DB 2; Length 236;

32	167	13.6	203	2	Q41340	Q41340 lycopersico
33	167	13.6	203	2	Q9M7P5	Q9m7p5 capsicum an
34	167	13.6	209	2	Q9U5G8	Q9u5g8 tetrahymena
35	166.5	13.5	202	2	Q08155	Q08155 pisum sativ
36	166.5	13.5	202	2	Q7DLK9	Q7dlk9 vicia faba
37	166.5	13.5	207	2	Q6AW64	Q6aw64 entamoeba h
38	166	13.5	212	2	Q7QWP4	Q7qwp4 giardia lam
39	166	13.5	212	2	Q9G079	Q9g079 giardia lam
40	165	13.4	202	2	Q9SXT5	Q9sxt5 cicer ariet
41	164.5	13.4	201	1	RB35 HUMAN	Q15286 homo sapien
42	164.5	13.4	201	2	Q6PHN9	Q6phn9 mus musculu
43	164.5	13.4	206	1	YPT5 CHLRE	Q39573 chlamydomon
44	164	13.3	192	2	Q9NDI3	Q9ndi3 entamoeba h
45	164	13.3	202	2	Q08153	Q08153 pisum sativ

ALIGNMENTS

RESULT 1

ID	Q8WUD3	PRELIMINARY;	PRT;	236 AA.
AC	Q8WUD3;			
DT	01-MAR-2002 (TrEMBLrel. 20, Created)			
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	RAB, member of RAS oncogene family-like 3.			
GN	Name=RABL3;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Testis;			
RX	MDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hoieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Testis;			
RA	Strausberg R.;			
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC020832; AAH20832.1; -			
DR	GO; GO:0003677; F:DNA binding; IEA.			
DR	GO; GO:0005525; P:small GTPase mediated signal transduction; IEA.			
DR	GO; GO:0007264; P:two-component signal transduction system (p. . .; IEA.			
DR	GO; GO:0000160; P:transcription factor binding; IEA.			
DR	PRINTS; PR00449; RASTRNSFRMNG.			
DR	PROSITE; PS00675; SIGMA54_INTERACT_1; 1.			
KW	GTP-binding.			
SQ	SEQUENCE 236 AA; 26308 MW; 13D9D19F8FD96710 CRC64;			

Best Local Similarity 99.2%; Pred. No. 2.7e-88;
Matches 234; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MASLDRVKVLVLGDSGVGKSSLVHLLCQNVLGPNPSWTGCVSDVRVHDYKEGTPEKTY 60
|||||
DB 1 MASLDRVKVLVLGDSGVGKSSLVHLLCQNVLGPNPSWTGCVSDVRVHDYKEGTPEKTC 60
|||||

QY 61 YIELWDVGGSGVGSASSVKSTRVFNYSVNGIIIFVHDLTNKKSSONLRWSLEALNRDLVP 120
|||||
DB 61 YIELWDVGGSGVGSASSVKSTRVFNYSVNGIIIFVHDLTNKKSSONLRWSLEALNRDLVP 120
|||||

QY 121 TGVLTNGDYGDOQFADNQIPLVIGTKLDQIHETKRHEVLRTAFLAEDFNPENLNLC 180
|||||

DB 121 TGVLTNGDYGDOQFADNQIPLVIGTKLDQIHETKRHEVLRTAFLAEDFNPENLNLC 180
|||||

QY 181 TNPRYLAAGSSNAVKLSRPFDKVIEKRYFLREGNQIPGFPDRKRFAGATLKSLHYD 236
|||||

DB 181 TNPRYLAAGSSNAVKLSRPFDKVIEKRYFLREGNQIPGFPDRKRFAGATLKSLHYD 236
|||||

RESULT 2

Q9D4V7 PRELIMINARY; PRT; 236 AA.

AC Q9D4V7; (T-EMBLrel. 17, Created)

DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)

DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)

DE Mus musculus adult male testis cDNA, RIKEN full-length enriched

DE library, clone:493053C05 product:hypothetical RAS small GTPases, Rab

DE subfamily/Arp/GTP-binding site motif A (P-loop)/Sigma-54 factor

DE interaction domain containing protein, full insert sequence (Rab13

DE protein);

GN Name=Rab13;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44(1999).

RN [2]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [3]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX The FANTOM Consortium.

RA The RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [4]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).

RN [5]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,

RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,

RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuiura S., Kawai J.,

RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format

RT sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).

RN [6]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,

RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,

RA Handgani T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,

RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,

RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,

RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,

RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,

RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,

RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,

RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,

RA Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [7]

SEQUENCE FROM N.A.

RC STRAIN=B5/SGFP transgenic ICR mice; TISSUE=Trophoblast Stem Cell;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalek U., Smailus D.B., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [8]

SEQUENCE FROM N.A.

RC STRAIN=B5/SGFP transgenic ICR mice; TISSUE=Trophoblast Stem Cell;

RA Strausberg R.;

RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK016099; BAB30113.1; -.

DR EMBL; BC050194; AAH50194.1; -.

DR MGD; MGI:1914907; Rab13.

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0005525; F:GTP binding; IEA.

DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.

DR GO; GO:0000160; P:two-component signal transduction system (p...; IEA.

DR InterPro; IPR001806; Ras trnsfrmg

DR InterPro; IPR002078; Sig54_interact.

DR Pfam; PF00071; Ras; 1.

DR PRINTS; PR00449; RASTRNSFRMG.

DR PROSITE; PS006675; SIGMAS4_INTERACT_1; 1.

KW GTP-binding; Hypothetical protein.

SQ SEQUENCE 236 AA; 26297 MW; DFCF72737C11944B CRC64;

Query Match 92.0%; Score 1133; DB 2; Length 236;
Best Local Similarity 92.4%; Pred. No. 1e-81;
Matches 218; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 MASLDRVKVLVLGDSGVGKSSLVHLLCQNVLGPNPSWTGCVSDVRVHDYKEGTPEKTY 60

```
Db 1 MASLDKRVKLVLDGSGVGKSSLVHLLCHNQVLGNPSTVCGSDVIRVHDYKEGTPPEKTY 60
Qy 61 YIELWDVGGSGVSGASSVKSTRVFNYSVNGIIFVHDLTNKKSSQNLRWLSLEALNRDLP 120
Db 61 YIELWDVGGSGVSGASSVKSTRVFNYSVNGIIFVHDLTNKKSSQNLRWLSLEALNRDLP 120
Qy 121 TGVLTNGDYDQGFADNQIPLVIGTKLQDIHETKRHEVLRTAFIAEDFNPEEINLDC 180
Db 121 TGVLTNGDYDQGFADNQIPLVIGTKLQDIHETKRHEVLRTAFIAEDFNPEEINLDC 180
Qy 181 TNPRYLAAGSSNAVKLSRPFDKVIEKRYFLREGNQIPGFPDRKRFAGTGLKSLHYD 236
Db 181 TNPRSSAAGSSNAVKLSRPFDKVIEKRYFFREGNQIPGFSDRKRFGGALKNFHCD 236

RESULT 3
ID Q8BMU2 PRELIMINARY; PRT; 236 AA.
AC Q8BMU2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:261031418 product:hypothetical RAS small
DE GTPases, Rab subfamily/ATP/GTP-binding site motif A (P-loop)/Sigma-54
DE factor interaction domain containing protein, full insert
DE sequence.
GN Name=Rabl3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuami T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
```

```
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
"RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuura T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK028195; BAC25804.1; -.
DR MGD; MGI:1914907; Rabl3.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR002078; Sig54_interact.
DR Pfam; PF00071; Ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR PROSITE; PS00675; SIGMAS4_INTERACT_1; 1.
KW GTP-binding; Hypothetical protein.
SQ SSQUENCE 236 AA; 26307 MW; 26307 MW; DFCET721377C20DB CRC64;

Query Match 92.0%; Score 1132; DB 2; Length 236;
Best Local Similarity 92.4%; Pred. No. 1-2e-81;
Matches 218; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

Qy 1 MASLDKRVKLVLDGSGVGKSSLVHLLCHNQVLGNPSTVCGSDVIRVHDYKEGTPPEKTY 60
Db 1 MASLDKRVKLVLDGSGVGKSSLVHLLCHNQVLGNPSTVCGSDVIRVHDYKEGTPPEKTY 60
Qy 61 YIELWDVGGSGVSGASSVKSTRVFNYSVNGIIFVHDLTNKKSSQNLRWLSLEALNRDLP 120
Db 61 YIELWDVGGSGVSGASSVKSTRVFNYSVNGIIFVHDLTNKKSSQNLRWLSLEALNRDLP 120
Qy 121 TGVLTNGDYDQGFADNQIPLVIGTKLQDIHETKRHEVLRTAFIAEDFNPEEINLDC 180
Db 121 TGVLTNGDYDQGFADNQIPLVIGTKLQDIHETKRHEVLRTAFIAEDFNPEEINLDC 180
Qy 181 TNPRYLAAGSSNAVKLSRPFDKVIEKRYFLREGNQIPGFPDRKRFAGTGLKSLHYD 236
Db 181 TNPRSSAAGSSNAVKLSRPFDKVIEKRYFFREGNQIPGFSDRKRFGGALKNFHCD 236

RESULT 4
ID Q6GPS4 PRELIMINARY; PRT; 235 AA.
AC Q6GPS4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC82648 protein.
GN Name=MGC82648;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RL MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
```

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnurch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT Initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Klein S., Gerhard D.S.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.
DR EMBL; BC073035; AAH73035.1; --
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR GO; GO:000160; P:two-component signal transduction system (p...); IEA.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR Pfam; PF00071; Ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
DR GTP-binding.
KW SEQUENCE 235 AA; 26180 MW; 9380A3E039C026CD CRC64;
SQ
Query Match 88.2%; Score 1085.5; DB 2; Length 235;
Best Local Similarity 90.3%; Pred. No. 5.8e-78;
Matches 213; Conservative 6; Mismatches 16; Indels 1; Gaps 1;
QY 1 MASLDRVKVLVLGDSGVGKSSLVHLLCQNVLGNSWTVGCSVDVRVHDYKEGTPEKTY 60
Db 1 MASLDRVKVLVLGDSGVGKSSLVHLLCQNVLGNSWTVGCSVDVRVHDYREGTPEKTY 60
QY 61 YIELWDVGGSGVSGASSVKSTRVFNYSVNGIIIFVHDLTNKSSQNLRRWSLEALNRDLVP 120
Db 61 YTELWDVGGSGVSGASSVKSTRVFNYSVNGIILVHDLTNKSSQNLRRWSLEALNRDLQP 120
QY 121 TGVLTNGDYGQDFQADNQLPLVIGTKLDQIHKHETKRHEVLITAFIAEDFNPENLDC 180
Db 121 MGVLTNGDYGDFQADNQLPLVIGTKLDQIPEAKRNEVLITAFIAEDFNAEINLDC 180
QY 181 TNPRYLAAGSSNAVKLSRPFDKVIKRYFLREGNQLPGPDKRFGAGTLKSLHYD 236
Db 181 TNTRCLAAGSSNAVKLSRPFDKVIKRY-PRGNLIPGFSDRKRRFGGNGFKSLHYD 235

RESULT 5

Q6TNS7
ID O6TNS7 PRELIMINARY; PRT; 233 AA.
AC Q6TNS7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 25-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Song H.D., Wu X.Y., Sun X.J., Zhou Y., Liu T.X., Deng M., Zhang G.W.,
RA Sheng Y., Chen Y., Ruan Z., Jiang C.L., Fan H.Y., Zou L.I.,
RA Kanki J.P., Look A.T., Chen Z.;
RA Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnurch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RA Strausberg R.;
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.
DR EMBL; AY391440; AAQ91252.1; --
DR EMBL; BC078191; AAH78191.1; --
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR GO; GO:000160; P:two-component signal transduction system (p...); IEA.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR002078; Sig54_interact.
DR Pfam; PF00071; Ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
DR GTP-binding; Hypothetical protein.
KW SEQUENCE 233 AA; 26113 MW; F1490193C884C086 CRC64;
SQ
Query Match 83.4%; Score 1026.5; DB 2; Length 233;
Best Local Similarity 83.5%; Pred. No. 2.7e-73;
Matches 197; Conservative 18; Mismatches 18; Indels 3; Gaps 1;
QY 1 MASLDRVKVLVLGDSGVGKSSLVHLLCQNVLGNSWTVGCSVDVRVHDYKEGTPEKTY 60
Db 1 MASLDRVKVLVLGDSGVGKSSLVHLLCQNVLGNSWTVGCSVDVRVHDYREGTPEKAF 60

```

QY 61 YIELWDVGGSGVSGASSVKSTRAVFNYSVNGIIFVHDLTNKSSQNLRRWSLEALNRDLVP 120
DB 61 YIELWDVGGSGVSGASSVKSTRAVFNYSVNGIIFVHDLTNKSSQNLRRWSLEALNRDLVP 120

QY 121 TGVLTNGDYYDOEFADNOLPLLVIGTKLQDIHETKRHEVLTRTAPLAEDFNPEINLDC 180
DB 121 TGIIVSNGDYDREQFAENAVPLLIGTKFKFOIPENKENDVLTRTAPLSEFNABEINLDC 180

QY 181 TNPRLAAGSNNAVKLSRPFKVKIEKRYFLREBGNQIFGPPDRKRFAGATLKSLHYD 236
DB 181 TNPRLAAGSNNAVKLSRPFKVKIEKRYFLRDPQMSQSFDRRRF---NFKSLHSD 233

RESULT 6
Q9D0M6 PRELIMINARY; PRT; 129 AA.
AC Q9D0M6
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2610002D21 product:hypothetical RAS small
DE GTPases, Rab subfamily/ATP/GTP-binding site motif A (P-loop)/Sigma-54
DE factor interaction domain containing protein, full insert
DE sequence.
GN Name=Rab13;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[2]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format

```

```

RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK011280; BAB27514.1; -
DR MGD; MGI:1914907; Rab13.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. .); IEA.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR002078; Sig54_interact.
DR Pfam; PF00071; Ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding; Hypothetical protein.
SQ SEQUENCE 129 AA; 14034 MW; 1B6E7723DF95499C CRC64;

Query Match 52.1%; Score 641; DB 2; Length 129;
Best Local Similarity 95.3%; Pred. No. 4.3e-43;
Matches 123; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MASLDVRKVLVLGDSGVGKSLVHLLCONQVLGNPSWTGCSVDVRVHDYKEGTPPEKTY 60
DB 1 MASLDVRKVLVLGDSGVGKSLVHLLCHNQVLGNPSWTGCSVDIRVHDYKEGTPPEKTY 60

QY 61 YIELWDVGGSGVSGASSVKSTRAVFNYSVNGIIFVHDLTNKSSQNLRRWSLEALNRDLVP 120
DB 61 YIELWDVGGSGVSGASSVKSTRAVFNYSVNGIILVHDLTNKSSQNLRRWSLEALNRDVP 120

QY 121 TGVLTNGD 129
DB 121 TGVLTNGD 129

RESULT 7
Q7Q5M9 PRELIMINARY; PRT; 236 AA.
AC Q7Q5M9
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP5724 (Fragment).
GN Name=agCG50989; ORFNames=ENSG000000010553;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
[1]
SEQUENCE FROM N.A.
STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008960; EAA10808.1; -
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.

```

DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR000345; CytC heme BS.
DR InterPro; IPR001806; Ras_Ernsfmg.
DR InterPro; IPR002078; Sig54_interact.
DR Pfam; PF00071; Ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding. 1 1
FT NON_TER 236 236
SQ SEQUENCE 236 AA; 26079 MW; 6FD2BA2D1D772A48 CRC64;

Query Match 48.2%; Score 593.5; DB 2; Length 236;
Best Local Similarity 53.9%; Pred.No.5.3e-39;
Matches 118; Conservative 41; Mismatches 45; Indels 15; Gaps 5;

QY 1 MASLDKVKVLVLDGSGVKSLSVHLHLCQNLGNPSWTGVCSDVVRVHYDKGTPEKTY 60
Db 15 MAALDKRVVLVCGDSGVKTSLSLTHLIANNEPLTSPGWTGVCSEVKLHVEGCTPAQNTF 74
QY 61 YIELDVGSGVGSASSVKSTRAVFNVSNGIIFVHDLTNKKSQNLRRNSLEALNRDLVP 120
Db 75 FVELDVGSGSI-----SHKNTRGVFNPTHGIIILVHDLTNKSKQNLRLMLAEILNKD--- 127
QY 121 TGVLVITNG---DYDQEQF-ADNQIPLLVITGTLKDOIHEKREHVLTRTA--FLAEDFNP 173
Db 128 -GDAUKGGDAIDMDPEQFGSTQIPILVITGTLKMDLDEGHGKIQITRTSAGSIASQCGA 186
QY 174 EEINLDTNTPRYLAGSSNAVKLSRFFDKVIEKRYFLRE 212
Db 187 DEILCLNCHESRLAAGTTDAVKLARFDKVKIERKYHNRD 225

RESULT 8
ID Q9VXA9 PRELIMINARY; PRT; 274 AA.
AC Q9VXA9; O8SZD5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE CG4789-PA (RE04047p).
GN ORFNames=CG4789;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA April J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flesler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwa Z.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M., M.G.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirekas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirekas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; AS003503; AAP48668.2; -.
DR ENBL; AY070945; AAL48567.1; -.
DR FlyBase; F8gn0030792; CG4789.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005525; F:GTP binding; IEA.

```
DR GO: 0007264; P: small GTPase mediated signal transduction; IEA.
DR GO: 0000160; P: two-component signal transduction system (p. . .; IEA.
DR InterPro: IPR001806; Ras trnsfrmg.
DR InterPro: IPR002078; Sig54_interact.
DR Pfam: PF00071; Ras; 1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding.
SQ SEQUENCE 274 AA; 30049 MW; 6A0E65C28875A7CC CRC64;

Query Match 40.8%; Score 502; DB 2; Length 274;
Best Local Similarity 43.2%; Pred. No. 1.1e-31;
Matches 114; Conservative 45; Mismatches 67; Indels 38; Gaps 8;

QY 1 MASLDKVLVLGDSGVGKSLVHLHCNQVLGNPSTVGCSDVDRVHDYKEGTPPEKTY 60
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MAMNRRVRIIVVGVDSGVGKSLVHLHCNQVLGNPSTVGCSDVDRVHDYKEGTPPEKTY 60
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 YIELMDVGVSGVSGVSSVKSSTRAVFNYSVNGIIFVHDLTNKSSQNLRRWSLEALNRD--- 117
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 FVELFDVGGSLNH-----KNTSRVFYAGIDGIIIVHDLTNKSSQNLRRWSLEALNRD--- 116
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 118 -----LVP-----TGVLVTNGD--VDQEQF-ADNQIPLLVIGTKLDQIHETK 156
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 117 TNKSGASMPSPSPSSSESTNLTGCHILFDMEEFLGATQTPIVLMGTGLDLDE-K 175
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 157 RHEVL--TRTAFLAEDFPNPEINLDCNTPRYLAAGSSNAVKLSRFPDKVIEKRYFLR--- 211
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 176 RHPKMGVKGPGGIADKCGAEIWLNCNRSLSLAAGTTDAVKLSRFPDRVIERKALRAAL 235
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 212 -----EGNQIPGFPDRKFRGAGTLK 231
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 236 AFGVSSNAVSPDRFRFGTSAK 259
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 9
Q948D2 PRELIMINARY; PRT; 354 AA.
AC Q948D2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative GTP-BINDING PROTEIN YPTV3.
GN Name=OSUNBA0013J21.21;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Saeki C., Henry D., Oates R., Simmons J.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079843; AL01181.1; -.
DR Gramene; Q948D2; -.
DR GO: 0003677; F: DNA binding; IEA.
DR GO: 0005525; P: GTP binding; IEA.
DR GO: 0007264; P: small GTPase mediated signal transduction; IEA.
DR GO: 0000160; P: two-component signal transduction system (p. . .; IEA.
DR InterPro: IPR001806; Ras trnsfrmg.
DR InterPro: IPR002078; Sig54_interact.
DR Pfam; PF00071; Ras; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding.
SQ SEQUENCE 354 AA; 38719 MW; DA15175B6A84A4A2 CRC64;

Query Match 22.4%; Score 276; DB 2; Length 354;
Best Local Similarity 29.8%; Pred. No. 1.2e-13;
Matches 75; Conservative 41; Mismatches 78; Indels 58; Gaps 9;

QY 6 RVKVLVLDGSGVSGKSLVHLHCNQVLGNPSTVGCSDVDR-----VHDYKEGTP- 55
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 27 QVRVLVVDGSGVSGKSLVHLHCNQVLGNPSTVGCSDVDR-----VHDYKEGTP- 86
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 56 -----BEKTYIYIELMDVGVSGVSGVSSVKSSTRAVFNYSVNGIIFVHDLTNKSS 103
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 87 SSSNSINSIKGDAERNFFVELMDVSGH-----ERYKECRSLFYSGINGVIFVYDLSQRKTK 142
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 104 QNLRRWSLEALNRDL--VPTGVLVTNGDYDQEQPADNQIPLLVIGTKLDQIHETKRHEVL 161
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 143 TNLNKWAVEVAESGTFPSAPLGGSGPG-----LPVPLVLTANKVDIAPRDKRVSS 193
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 162 TRTAFLAEDF-----NPPEINLDCNTPRYLAAGSSNAVKLSR-----PFDKVIE 205
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 10
Q7X985 PRELIMINARY; PRT; 354 AA.
AC Q7X985;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative GTP-BINDING PROTEIN YPTV3.
GN ORFNames=OSUNBA0013J21.21;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA The Rice Chromosome 10 Sequencing Consortium;
RA "In-depth view of structure, activity, and evolution of rice
RA chromosome 10.";
RL Science 300:1566-1569 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017054; AAP52031.1; -.
DR Gramene; Q7X985; -.
DR GO: 0003677; F: DNA binding; IEA.
DR GO: 0005525; P: GTP binding; IEA.
DR GO: 0007264; P: small GTPase mediated signal transduction; IEA.
DR GO: 0000160; P: two-component signal transduction system (p. . .; IEA.
DR InterPro: IPR001806; Ras trnsfrmg.
DR InterPro: IPR002078; Sig54_interact.
DR Pfam; PF00071; Ras; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding.
SQ SEQUENCE 354 AA; 38719 MW; DA15175B6A84A4A2 CRC64;

Query Match 22.4%; Score 276; DB 2; Length 354;
Best Local Similarity 29.8%; Pred. No. 1.2e-13;
Matches 75; Conservative 41; Mismatches 78; Indels 58; Gaps 9;

QY 6 RVKVLVLDGSGVSGKSLVHLHCNQVLGNPSTVGCSDVDR-----VHDYKEGTP- 55
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 27 QVRVLVVDGSGVSGKSLVHLHCNQVLGNPSTVGCSDVDR-----VHDYKEGTP- 86
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 56 -----BEKTYIYIELMDVGVSGVSGVSSVKSSTRAVFNYSVNGIIFVHDLTNKSS 103
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 87 SSSNSINSIKGDAERNFFVELMDVSGH-----ERYKECRSLFYSGINGVIFVYDLSQRKTK 142
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 104 QNLRRWSLEALNRDL--VPTGVLVTNGDYDQEQPADNQIPLLVIGTKLDQIHETKRHEVL 161
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 143 TNLNKWAVEVAESGTFPSAPLGGSGPG-----LPVPLVLTANKVDIAPRDKRVSS 193
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 162 TRTAFLAEDF-----NPPEINLDCNTPRYLAAGSSNAVKLSR-----PFDKVIE 205
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

Db 194 GNLVDVAFQWVEKQGLLSESLPLAESPPG--NSGLLTAAKVAKYDKALVKFFRMLIR 251
 Qy 206 KRYFLREGNQIP 217
 Db 252 RRYF---SNELP 260

RESULT 11
 Y813 ARATH
 ID Y813 ARATH STANDARD; PRT; 342 AA.
 AC Q9CSJ9; Q9LV96;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hypothetical GTP-binding protein At5g64813.
 GN OrderedLocustNames=At5g64813; ORFNames=MXK3.3;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RA "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:31-63(2000).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Akarawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Heuan W.W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinzaki K., Davis R.W., Theologis A., Becker J.R.;
 RA "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome.";
 RL Science 302:942-946(2003).
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to erroneous
 CC gene model prediction. The predicted gene should be splitted into
 CC 3 genes: At5g64810, At5g64813 and At5g64816.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on it
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercia
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB019236; BAA97293.1; ALT_SEQ.
 CC EMBL; AF360203; AAK25913.1; -.
 CC EMBL; AY040050; AAK64108.1; -.
 CC InterPro; IPR001806; Ras_trnsfrmg.
 CC Pfam; PF00071; Ras; 1.
 CC PRINTS; PR00449; RASTRNSFRMG.
 CC GTP-binding; Hypothetical protein.
 CC DOMAIN 71 75 Poly-Ser.
 CC FT NP_BIND 29 36 GTP (Potential).
 CC FT NP_BIND 90 94 GTP (Potential).
 CC FT

```
Qy 52 EGPTEPKYIYELWDVGGSSGVSSSVKSTRVFNVSNGIIFVHDLTNKSSQNLRRW-- 109
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 219 AGTPEQTELELWDIGGVAHQAQAQ-----VFPEGAAGVAILVHDLTNKRSEENLATWLT 274

Qy 110 SLEALNRDLVPTGVLTNGDYDQEPADNPDIPLLVIGTKLDQI-HE-----TKBEVLVTRT 164
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 275 MLGDKPRGAAPKSKDPAVALKVDIESCNPLVLTGKADLVPHKQFVSDYDRYDGLGNA 334

Qy 165 AFLAEDFNPEINLDCNTNPRYLAAGSSNAVKLSRFFDKVIEK-----FYFLREG 213
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 335 -----NSAISVWKPNNRRKNSKKPMKIGR--KLIVKEKVKVFGQYLRG 378

RESULT 13
Q9FIB5 PRELIMINARY; PRT; 431 AA.
AC Q9FIB5;
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone:MYH9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RX MEDLINE=99156233; PubMed=10048488;
RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
RT Sequence features of the regions of 1,081,958 bp covered by seventeen
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:379-391(1998).
DR ENBL; AB016893; BAB09412.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003822; PAH.
SQ SEQUENCE 431 AA; 48540 MW; PFDJ3C8C7D1D1167 CRC64;

Query Match 18.2%; Score 224; DB 2; Length 431;
Best Local Similarity 27.5%; Pred. No. 2e-09;
Matches 66; Conservative 40; Mismatches 66; Indels 68; Gaps 11;

Qy 6 RVKVLVLDGSGVSKSLVHLLCQNVLGNSWTGCSVDVVRVHDYKEGT-----PER 57
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 22 QIRVLVVGDSG-----SSIV-----RPSQTIGCTGVGKHLTYASPASSSSIIKGDSE 68

Qy 58 KTYIYELWDVGGSSGVSSSVKSTRVFNVSNGIIFVHDLTNKSSQNLRRWSLEALNRD 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 69 RDPFVELWDVSGH-----ERYKDCSLFYSQINGVIFVHDLISQRTKTNLQNA----- 117

Qy 118 LVPTGVLTNGDYDQEPADN---QIPLLVIGTKLD-----QIHETKRHEVL 161
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 118 ----GEVSVTGFESAPLSGCGPLPVYIVIGNKADIAAKGTGNGSGNLVDAARHWE 173

Qy 162 TRTAF-----LAEDFNPEINLDCNTNPRYLAAGSSNAVK--LSRFFDKVIEKRYFLRE 212
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 174 KQGLPHSDPLPSSEF-PSNVGL-----IMAAKEARYDKEALTQIFHMLIRRRYFSD 226

RESULT 14
Q9VQG6 PRELIMINARY; PRT; 242 AA.
AC Q9VQG6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CG15399-PA.
GN ORFNames=CG15399;
OS Drosophila melanogaster (Fruit fly).
```

```
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley B.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Liang Y., Lin X.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Swirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Swirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Swirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
```

DR InterPro; IPR005225; Small_GTP.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 23:29:44 ; Search time 29 Seconds
(without alignments)
607.489 Million cell updates/sec

Title: US-09-945-173-2

Perfect score: 1231

Sequence: 1 MASLDRVKVLVLDGSGVGKS.....PGFPDRKRFAGTILKSLHYD 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*

1: /cgm2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgm2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgm2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgm2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgm2_6/ptodata/1/iaa/PTUS_COMB.pep.*

6: /cgm2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	450	36.6	325	4	US-09-270-767-45478
2	166.5	13.5	228	4	US-09-270-767-46812
3	164.5	13.4	228	4	US-09-949-016-9506
4	163	13.2	201	2	US-08-531-525-13
5	163	13.2	201	2	US-08-718-270A-13
6	156.5	12.7	201	2	US-08-916-901-8
7	156.5	12.7	201	3	US-09-154-602-8
8	156	12.7	228	4	US-09-270-767-46283
9	155.5	12.6	203	4	US-08-255-920A-12
10	154.5	12.6	201	2	US-08-916-901-3
11	154.5	12.6	201	3	US-09-154-602-3
12	154.5	12.6	202	2	US-08-531-525-14
13	154.5	12.6	202	2	US-08-718-270A-14
14	154.5	12.6	205	4	US-09-709-103-49
15	154.5	12.6	205	4	US-09-439-410A-49
16	154.5	12.6	237	4	US-09-949-016-7643
17	154	12.5	197	4	US-09-270-767-45823
18	153	12.4	206	4	US-09-828-310-13
19	152.5	12.4	215	2	US-08-531-525-10
20	152.5	12.4	215	2	US-08-718-270A-10
21	150.5	12.2	238	4	US-09-248-796A-19962
22	149.5	12.1	202	2	US-08-531-525-24
23	149.5	12.1	202	2	US-08-718-270A-24
24	149	12.1	207	2	US-08-824-873-4
25	149	12.1	207	3	US-09-156-184-4
26	148	12.0	227	4	US-09-270-767-43285
27	148	12.0	227	4	US-09-248-796A-20291

28 147.5 12.0 342 4 US-09-949-016-10797 Sequence 10797, A
29 146.5 11.9 131 4 US-09-270-767-58626 Sequence 58626, A
30 146.5 11.9 214 4 US-09-270-767-33012 Sequence 33012, A
31 146.5 11.9 214 4 US-09-270-767-48229 Sequence 48229, A
32 145 11.8 369 4 US-09-949-016-7355 Sequence 7355, A
33 144.5 11.7 191 2 US-08-531-525-26 Sequence 26, Appl
34 144.5 11.7 191 2 US-08-718-270A-26 Sequence 26, Appl
35 143.5 11.7 203 4 US-09-949-016-9379 Sequence 9379, A
36 143.5 11.7 214 3 US-08-741-411-11 Sequence 11, Appl
37 142.5 11.6 205 2 US-08-531-525-25 Sequence 25, Appl
38 142.5 11.6 205 2 US-08-718-270A-25 Sequence 25, Appl
39 142.5 11.6 213 4 US-09-248-796A-20293 Sequence 20293, A
40 142 11.5 128 4 US-09-513-999C-5775 Sequence 5775, A
41 142 11.5 194 3 US-08-741-411-9 Sequence 9, Appl
42 142 11.5 203 4 US-09-949-016-6440 Sequence 6440, A
43 142 11.5 213 2 US-08-531-525-12 Sequence 12, Appl
44 142 11.5 213 2 US-08-718-270A-12 Sequence 12, Appl
45 142 11.5 216 4 US-09-949-016-7056 Sequence 7056, A

ALIGNMENTS

RESULT 1

US-09-270-767-45478

; Sequence 45478, Application US/09270767

; Patent NO. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270.767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 45478

; LENGTH: 325

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-45478

Query Match 36.6%; Score 450; DB 4; Length 325;
Best Local Similarity 42.1%; Pred. No. 7.1e-42;
Matches 104; Conservative 40; Mismatches 65; Indels 38; Gaps 8;

Qy	18	GKSLVHLLCQNVLGNPSTVCGSDVRVDYKEGTPEETYYIELWDVCGSVGSASSV	77
Db	70	GKTSLSLTHLTHNEALIRPGWTVGCNIQVQMHPPREGTARECPYFVELFDVGGSLNH	125
Qy	78	KSTRVFPYNSVNGIIFVHDLTNKKSSQNLRRWSLEALNRD	120
Db	126	KNTSVEFYAGIDGILLVHDLITNAKSQRLDYLVEIVNKEGKDTNKGASMPSPSPPL	185
Qy	121	-----TGVLTNGD--YDOEQF-ADNQIPLVIGTKLDOIHTKREHVL--TRTAPLADF	171
Db	186	SSFTDNLGTDGHLFDMEEFGLGATQPIILVMGTKLLDDE-KRHPKMGVKKPGGIADKC	244
Qy	172	NPEHINDCTNPRYLAAGSSNAVKLSRFPDKVIEKRFILR	224
Db	245	GAEIWLNCNRSRLAAGTTDAVKLSRFPDRVIENKRALRAALAFGVSSNNAVPPDRRR	304
Qy	225	FGAGTLK 231	
Db	305	FGPTSAK 311	

RESULT 2

US-09-270-767-46812

; Sequence 46812, Application US/09270767

; Patent NO. 6703491

; GENERAL INFORMATION:

```

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46812
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46812

Query Match      13.5%; Score 166.5; DB 4; Length 227;
Best Local Similarity 31.5%; Pred. No. 3e-10;
Matches 40; Conservative 25; Mismatches 53; Indels .9; Gaps 2;

Qy      8 KVLVLGDSGVGKSSLVHLLCONVLGNPSWTVCSDVVRVHDYKEGTPEKTYTIELMDV 67
Db      36 KLLIIGDTGVGKSSILLIRFSDTFSGSYITITGVDFKIRTVDI-----EGMRVKLIQIWD 90
Qy      68 GGSVGSASSVSKTRAFVNSVNGIIFVHDLTNKSSQNLKRWSLKALNRDLVPTGVLVTN 127
Db      91 AGQ-----ERFRITITTYRGTGHVIVVDVTVNGSFAVNRVWLEBIQNNCDVVKVLGVN 146
Qy      128 GYDQEQ 134
Db      147 KNDDPDR 153

RESULT 3
US-09-949-016-9506
; Sequence 9506, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9506
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9506

Query Match      13.4%; Score 164.5; DB 4; Length 228;
Best Local Similarity 25.9%; Pred. No. 5.1e-10;
Matches 53; Conservative 36; Mismatches 63; Indels 53; Gaps 9;

Qy      8 KVLVLGDSGVGKSSLVHLLCONVLGNPSWTVCSDVVRVHDYKEGTPE---EKTYYIEL 64
Db      37 KLLIIGDSGVGKSSILLIRFADNTFSGSYITIGV-----DFKIRTVINGEKV-KLQI 88
Qy      65 WDVGSGVGSASVSKSTRAFVNSVNGIIFVHDLTNKSSQNLKRWSLKALNRDLVPTGVL 124
Db      89 WDTAGQ---ERFRITITTYRGTGHVIVVDVTVNGSFAVNRVWLEBIQNNCDVVKVLGVN 137
Qy      125 VTNGDYDOEQADNQIPLLVITGTKLDQIHETKREHVLTFETAF-----LAED 170
Db      138 -----DDVCKRILVGNKND---DPERKVVETEDAYKFGQMGICQLFETSAREN 181
Qy      171 FNPPEINLDCNTPRYLAAGSSNAVK 195

```

US-08-718-270A-13

Query Match 13.2%; Score 163; DB 2; Length 201;
Best Local Similarity 25.9%; Pred. No. 6.1e-10;
Matches 51; Conservative 42; Mismatches 80; Indels 24; Gaps 6;
Qy 8 KVLVLGDSGVKSSLVHLLCQNVLGNSPWTGCVSDVRVHDYKEGTPPEKTYIYELWDV 67
Db 10 KLLIGDSGVKSCLLRFADDTTYSYISTIGVDFKIRTEL-----DGKTIKLIQIWD 64
Qy 68 GGSVGSASSVKSTRVAFYNSVNGIIFVHDLTNKSSQNLRLRW-----SLEALNRDLVP 120
Db 65 AGQ-----ERFTTSSYRGAGHIIIVYDVTQDSFNNVKQLQEI-----ADRYASENVKLLVG 120
Qy 121 TGVLVWTG---DYDQEQFADN-QIPLLV-----GKLDQIHETKREHVLTRTAFLAEDFN 172
Db 121 NKCDLTTKVVYDTTKGFADSLGIFPFLTSKATNATVEQAFMTMAAEIKKMGFGATSGG 180
Qy 173 PEINLDCNPRYLAAG 189
Db 181 SEKSNVIQSTPVKSSG 197

US-08-718-270A-13

Query Match 13.2%; Score 163; DB 2; Length 201;
Best Local Similarity 25.9%; Pred. No. 6.1e-10;
Matches 51; Conservative 42; Mismatches 80; Indels 24; Gaps 6;
Qy 8 KVLVLGDSGVKSSLVHLLCQNVLGNSPWTGCVSDVRVHDYKEGTPPEKTYIYELWDV 67
Db 10 KLLIGDSGVKSCLLRFADDTTYSYISTIGVDFKIRTEL-----DGKTIKLIQIWD 64
Qy 68 GGSVGSASSVKSTRVAFYNSVNGIIFVHDLTNKSSQNLRLRW-----SLEALNRDLVP 120
Db 65 AGQ-----ERFTTSSYRGAGHIIIVYDVTQDSFNNVKQLQEI-----ADRYASENVKLLVG 120
Qy 121 TGVLVWTG---DYDQEQFADN-QIPLLV-----GKLDQIHETKREHVLTRTAFLAEDFN 172
Db 121 NKCDLTTKVVYDTTKGFADSLGIFPFLTSKATNATVEQAFMTMAAEIKKMGFGATSGG 180
Qy 173 PEINLDCNPRYLAAG 189
Db 181 SEKSNVIQSTPVKSSG 197

US-08-718-270A-13

Query Match 13.2%; Score 163; DB 2; Length 201;
Best Local Similarity 25.9%; Pred. No. 6.1e-10;
Matches 51; Conservative 42; Mismatches 80; Indels 24; Gaps 6;
Qy 8 KVLVLGDSGVKSSLVHLLCQNVLGNSPWTGCVSDVRVHDYKEGTPPEKTYIYELWDV 67
Db 10 KLLIGDSGVKSCLLRFADDTTYSYISTIGVDFKIRTEL-----DGKTIKLIQIWD 64
Qy 68 GGSVGSASSVKSTRVAFYNSVNGIIFVHDLTNKSSQNLRLRW-----SLEALNRDLVP 120
Db 65 AGQ-----ERFTTSSYRGAGHIIIVYDVTQDSFNNVKQLQEI-----ADRYASENVKLLVG 120
Qy 121 TGVLVWTG---DYDQEQFADN-QIPLLV-----GKLDQIHETKREHVLTRTAFLAEDFN 172
Db 121 NKCDLTTKVVYDTTKGFADSLGIFPFLTSKATNATVEQAFMTMAAEIKKMGFGATSGG 180
Qy 173 PEINLDCNPRYLAAG 189
Db 181 SEKSNVIQSTPVKSSG 197

RESULT 6
US-08-916-901-8
; Sequence 8, Application US/08916901
; Patent No. 5892012
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/08/916,901
; APPLICATION NUMBER: US/08/916,901
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 57006
; US-08-916-901-8

Query Match 12.7%; Score 156.5; DB 2; Length 201;
Best Local Similarity 24.1%; Pred. No. 3.3e-09;
Matches 49; Conservative 42; Mismatches 69; Indels 43; Gaps 8;
Qy 8 KVLVLGDSGVKSSLVHLLCQNVLGNSPWTGCVSDVRVHDYKEGTPPEKTYIYELWDV 67
Db 10 KLLIGDSGVKSCLLRFADDTTYSYISTIGVDFKIRTEL-----DGKTIKLIQIWD 64
Qy 68 GGSVGSASSVKSTRVAFYNSVNGIIFVHDLTNKSSQNLRLRW-----SLEALNRDLVP 127
Db 65 AGQ-----ERFTTSSYRGAGHIIIVYDVTQDSFNNVKQLQEI-----ADRYASENVKLLVG 106
Qy 128 GDYDQEQFADN-QIPLLV-----GKLDQIHETKREHVLTRTAFLAEDFNPEINLDCNPRYLA 187
Db 107 -----ADRYASENVKLLVGNSDL--TTKK--VVDNTT--AKEP-----ADSLGVFPFLE 149
Qy 188 AGSSNAVKLSRFPDKV---IEKR 207
Db 150 TSAKNATNVEQAFMTMAAEIKKR 172

RESULT 7
US-09-154-602-8
; Sequence 8, Application US/09154602
; Patent No. 6300472

```

; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti C.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,602
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/916,901
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 57006
; US-09-154-602-8

Query Match 12.7%; Score 156.5; DB 3; Length 201;
Best Local Similarity 24.1%; Pred. No. 3.3e-09;
Matches 49; Conservative 42; Mismatches 69; Indels 43; Gaps 8;

QY 8 KVLVLGDSGVGKSSLVHLLCQNV--LGNPSWTGCSVDVRVHDYKEGTPBEKTYIELWDV 67
Db 10 KLLIGDSGVGKSCLLLRFPADDTYTESYSTIGVDFKIRTIEL-----DGKTIKLIQWDT 64

QY 68 GGSVGSASSVKSTRVFNYSNGIIFVHDLTNKSSQNLRRWSLEALNRDLVPTGVLVTN 127
Db 65 AGQ-----ERFTTSSYRGAGHIIIVVDVTDQESYANVKQMLQEI----- 106

QY 128 GDYDQEQFADNQIPLLVIGTKLQIHEHVKHEVLTAFIAEDFNPEEINLDCNPRYLA 187
Db 107 -----DRIASENVKLLVGNKSL--TTKK--VVDNNT--AKEF-----ADSLGVPFLE 149

QY 188 AGSSNAVKLSRFPDKV---IEKR 207
Db 150 TSAKNATNVEQAFTWAAEIKR 172

RESULT 8
US-09-270-767-46283
; Sequence 46283, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46283
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-09-270-767-46283

Query Match 12.7%; Score 156; DB 4; Length 228;
Best Local Similarity 27.5%; Pred. No. 4.6e-09;
Matches 44; Conservative 30; Mismatches 56; Indels 30; Gaps 5;

QY 8 KVLVLGDSGVGKSSLVHLLCQNV--LGNPSWTGCSVDVRVHDYKEGTPBEKTYIELW 65
Db 21 RLILGDSVTGKSSLLKFTDGRFAELSDP--TVGVDFARLIEMKDGQIK-----LQLW 74

QY 66 DVGSGVGSASSVKSTRVFNYSNGIIFVHDLTNKSSQNLRRWSLEALNRDLVPTGVLV 125
Db 75 DTAGQERFRTSITKS-----YRNSVGVLLVYDISNHPASPEHPLMMWEA----- 118

QY 126 TNGDYDQEQFADNQIPLLVIGTKLQIHEHVKHEVLTRTA 165
Db 119 -----QRHIEPRPVPFALVGCKLDLINAGHREVTTEA 152

RESULT 9
US-09-255-920A-12
; Sequence 12, Application US/09255920A
; Patent No. 6623980
; GENERAL INFORMATION:
; APPLICANT: Fisher, Joseph
; APPLICANT: Lorens, James
; APPLICANT: Anderson, David
; APPLICANT: Luo, Ying
; APPLICANT: Huang, Betty
; APPLICANT: Shen, Mary
; TITLE OF INVENTION: EXO1 and EXO2, EXOCYTOTIC PROTEINS
; FILE REFERENCE: A65905-1/DJB/RMS
; CURRENT APPLICATION NUMBER: US/09/255,920A
; CURRENT FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: 60/075,534
; PRIOR FILING DATE: 1998-02-23
; PRIOR APPLICATION NUMBER: 60/086,650
; PRIOR FILING DATE: 1998-05-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (34)
; OTHER INFORMATION: The xaa at position 34 represents an unknown amino
; OTHER INFORMATION: acid.
; US-09-255-920A-12

Query Match 12.6%; Score 155.5; DB 4; Length 203;
Best Local Similarity 23.8%; Pred. No. 4.3e-09;
Matches 51; Conservative 39; Mismatches 69; Indels 55; Gaps 9;

QY 8 KVLVLGDSGVGKSSLVHLLCQNVGNPSW--TVGCSVDVRVHDYKEGTPBEKTYIELW 65
Db 10 KLLIGDSGVGKTCVLFREFSD--AFXNSTFISTIGIDFKIRTIEL-----DGKRIKLIW 63

QY 66 DVGSGVGSASSVKSTRVFNYSNGIIFVHDLTNKSSQNLRRWSLEALNRDLVPTGVLV 125
Db 64 DTAGQ-----ERFTTITAYRGAMGIMLVYDITNEKSPFNIRNWI-----RNI----- 107

QY 126 TNGDYDQEQFADNQIPLLVIGTKLQIHEHVKHEVLTRTAFLAEDFN-----PEEIN 177
Db 108 -----EEHASADVEMKILGNKCDV--NDKRVSKERGEKLDYGIKFMETSAKANIN 158
```


; GENERAL INFORMATION:
 ; APPLICANT: Hlavka, Joseph J.
 ; APPLICANT: Pincus, Matthew R.
 ; APPLICANT: No. 58406831e, John F.
 ; APPLICANT: Abajian, Henry B.
 ; APPLICANT: Kende, Andrew S.
 ; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; STREET: 5370 Manhattan Circle, Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: US
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/531,525
 ; FILING DATE: 21-SEP-1995
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ferber, Donna M.
 ; REGISTRATION NUMBER: 33,878
 ; REFERENCE/DOCKET NUMBER: 37-94
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (303) 499-8080
 ; TELEFAX: (303) 499-8089
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 202 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Lymnea stagnalis
 ; US-08-531-525-14

Query Match 12.6%; Score 154.5; DB 2; Length 202;
 Best Local Similarity 22.8%; Pred. No. 5.6e-09;
 Matches 44; Conservative 38; Mismatches 70; Indels 41; Gaps 5;
 QY 8 KVLVLGDSGVGKSLVHLLCQNVLGNSPWTVCSDVVRVHDYKEGTPPEKTYIYELWDV 67
 Db 13 KLLLLGDSGVGKSLVHLLCQNVLGNSPWTVCSDVVRVHDYKEGTPPEKTYIYELWDV 67
 QY 68 GGSVGSASSVKSTRVFNYSVNGIIFVHDLTNKSSQNLRWSLEALNRDLVPTGLVTN 127
 Db 68 AGQ----ERPTITSSYRGAGHIIIVYDVTDOESFNVKQWLQEI----- 109
 QY 128 GDYDQBFADNQIPLVIGTKLDQIHETKREHVLTRTAFIAEDFNPEEINLDCNRYLA 187
 Db 110 -----DRYASENVNKLGVNKSDDL--TTKK-----VDFTTAKEYADQLGIPFLE 151
 QY 188 AGSSNAVKLSRFF 200
 Db 152 TSAKNATNVEQAF 164

RESULT 13
 US-08-718-270A-14
 ; Sequence 14, Application US/08718270A
 ; Patent No. 5910478
 ; GENERAL INFORMATION:
 ; APPLICANT: Hlavka, Joseph J.
 ; APPLICANT: Pincus, Matthew R.
 ; APPLICANT: No. 59104781e, John F.

; APPLICANT: Abajian, Henry B.
 ; APPLICANT: Kende, Andrew S.
 ; TITLE OF INVENTION: Peptidomimetics Inhibiting
 ; TITLE OF INVENTION: the Oncogenic Action of P21 Ras
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 ; STREET: 5370 Manhattan Circle, Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: US
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/718,270A
 ; FILING DATE: 20-SEP-1996
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/531,525
 ; FILING DATE: 21-SEP-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/004,091
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ferber, Donna M.
 ; REGISTRATION NUMBER: 33,878
 ; REFERENCE/DOCKET NUMBER: 78-95
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (303) 499-8080
 ; TELEFAX: (303) 499-8089
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 202 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Lymnea stagnalis
 ; US-08-718-270A-14

Query Match 12.6%; Score 154.5; DB 2; Length 202;
 Best Local Similarity 22.8%; Pred. No. 5.6e-09;
 Matches 44; Conservative 38; Mismatches 70; Indels 41; Gaps 5;
 QY 8 KVLVLGDSGVGKSLVHLLCQNVLGNSPWTVCSDVVRVHDYKEGTPPEKTYIYELWDV 67
 Db 13 KLLLLGDSGVGKSLVHLLCQNVLGNSPWTVCSDVVRVHDYKEGTPPEKTYIYELWDV 67
 QY 68 GGSVGSASSVKSTRVFNYSVNGIIFVHDLTNKSSQNLRWSLEALNRDLVPTGLVTN 127
 Db 68 AGQ----ERPTITSSYRGAGHIIIVYDVTDOESFNVKQWLQEI----- 109
 QY 128 GDYDQBFADNQIPLVIGTKLDQIHETKREHVLTRTAFIAEDFNPEEINLDCNRYLA 187
 Db 110 -----DRYASENVNKLGVNKSDDL--TTKK-----VDFTTAKEYADQLGIPFLE 151
 QY 188 AGSSNAVKLSRFF 200
 Db 152 TSAKNATNVEQAF 164

RESULT 14
 US-09-709-103-49
 ; Sequence 49, Application US/09709103
 ; Patent No. 6733991
 ; GENERAL INFORMATION:
 ; APPLICANT: Cismowski, Mary

Search completed: July 1, 2005, 00:02:36
Job time : 31 secs

```

RESULT 15
US-09-439-410A-49
; Sequence 49, Application US/09439410A
; Patent No. 6746852
; GENERAL INFORMATION:
; APPLICANT: Cismowski, Mary
; APPLICANT: Duzic, Emir
; TITLE OF INVENTION: AGS PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: 1919/60388-B
; CURRENT APPLICATION NUMBER: US/09/439,410A
; CURRENT FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-439-410A-49

```

Query Match	12.6%;	Score 154.5;	DB 4;	Length 205;
Best Local Similarity	22.7%;	Pred. No. 5.7e-09;		
Matches	46;	Conservative 41;	Mismatches 73;	Indels 43; Gaps 6;
QY	8	KVLVLGDSGVGKSGSLVHLLQCONVLGNPMSVTGCSVDVRVHDYKEGTPPEETKYVIELWDV	67	
DB	13	KLLLLGDSGVGKSCLLLPADDTYTESYISTIGVDFKIRIEL-----DGKTIKQLQIWDT	67	
QY	68	GGSGVGSASSVSTRAPVFNYSVNGIIFVHDLTNKKSQNLRWSLEALNRDLVPPTGVILVTN	127	
DB	68	AGQ-----ERFRTITSSYRGAHGIIIVVDVTDQESFNNVKQWLOEI-----	109	
QY	128	GDYDQEQFADNQPLLVIGTGKLDQIHETKREHLVTRTFLAEDPNPEINLDCNPRYLA	187	
DB	110	-----DRIASENVNKLIVGNKCDL-----TTKK-----VDYTTAKEPADSLGIFPLE	152	
QY	188	AGSSNAVKLSRFPDKV---IEKR	207	
DB	153	TSAKNATNVEQSFMTMAAEIKR	175	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 13:28:23 ; Search time 2650.06 Seconds
(without alignments)
10212.469 Million cell updates/sec

Title: US-09-945-173-3
Perfect score: 711
Sequence: 1 atgggcctcctgagctgggt.....agagccttcattgactga 711

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	686.2	96.5	918	4	BI862084
2	684	96.2	718	5	BX099226
3	682.2	95.9	720	4	BG501635
4	681.4	95.8	732	4	BI333788
5	680.6	95.7	699	7	CN361072
6	679.4	95.6	720	5	CN361073
7	674.8	94.9	815	5	BQ215951
8	668	94.3	1867	3	CR619977
9	663.6	93.3	797	1	AL516309
10	654.4	92.0	989	4	BM451532
11	650.2	91.4	902	4	BG258348
12	650	91.4	855	7	CK455329
13	644	90.6	785	4	BI759699
14	636.4	89.5	649	4	BG025782
15	628	88.3	959	4	BG031602
16	613.6	86.3	696	7	CN794203
17	611	85.9	986	4	BI334986
18	589.8	83.0	920	1	AA203636
19	588.8	82.8	602	4	BG613092
20	587.8	82.7	993	3	AK016099
21	586.2	82.4	2013	3	AK028195
22	583	82.0	584	7	CN361075
23	581	81.7	875	7	CK456385
24	576.8	81.1	2006	3	AK028269

25	573.8	80.7	706	7	CF730316
26	562.6	79.1	665	4	BG080166
27	559	78.6	786	2	BF106997
28	558	78.5	827	2	BF571796
29	548.8	77.2	722	2	BB610950
30	542.4	76.3	743	2	BB611463
31	540	75.9	551	1	AA662211
32	534.4	75.2	810	6	CD512097
33	533.4	75.0	536	7	CN361079
34	531	74.7	573	2	AW373567
35	525.6	73.9	1053	4	BM015053
36	514	72.3	526	5	BU431648
37	511	71.9	518	4	BM147898
38	507	71.3	563	2	BF208017
39	503.6	70.8	826	1	AJ456809
40	497.6	70.0	798	6	CD298863
41	491.6	69.1	598	5	BU706219
42	483.4	68.0	494	7	CN361078
43	481.2	67.7	800	5	BU460217
44	478.2	67.3	573	6	CA542843
45	477.4	67.1	536	4	BG610366

ALIGNMENTS

RESULT 1
LOCUS BI862084 918 bp mRNA linear EST 10-OCT-2001
DEFINITION G0339070F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5399969 5', mRNA sequence.
ACCESSION BI862084
VERSION BI862084.1 GI:16002831
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 918)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@nsl.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM12019 row: 1 column: 18
High quality sequence stop: 833.
Location/Qualifiers
1. 918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5399969"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

FEATURES

source
Query Match 96.5%; Score 686.2; DB 4; Length 918;
Best Local Similarity 99.3%; Pred. No. 4e-189;
Matches 700; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

```
QY 9 CTTGGATCGGTGAAGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTAGT 68
Db |||
QY 7 CTTGGATCGGTGAAGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTAGT 66
Db |||
QY 69 CCATCTCTTATGCAAAATCAAGTGTGGGAAATCCATCATGAGCTGTGGGCTGCTCAGT 128
Db |||
QY 67 CCATCTCTTATGCAAAATCAAGTGTGGGAAATCCATCATGAGCTGTGGGCTGCTCAGT 126
Db |||
QY 129 GGNATGTCAGATTCATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 188
Db |||
QY 127 GGNATGTCAGATTCATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 186
Db |||
QY 189 ATTATGGGATGTTGGAGGCTCTGTGGGAGTGCAGCAGCGTGAAGAGACAGAGAGAGT 248
Db |||
QY 187 ATTATGGGATGTTGGAGGCTCTGTGGGAGTGCAGCAGCGTGAAGAGACAGAGAGAGT 246
Db |||
QY 249 ATTCTCAACTCGTAAATGGTATTAATTTTGTGACAGACTTAACAAATAAGAGAGTCTTC 308
Db |||
QY 247 ATTCTCAACTCGTAAATGGTATTAATTTTGTGACAGACTTAACAAATAAGAGAGTCTTC 306
Db |||
QY 309 CCAAAACTTGGCGTGTGCTCATGGAAGCTCTCAACAGGATTTGGTCCCACTGGAGT 368
Db |||
QY 307 CCAAAACTTGGCGTGTGCTCATGGAAGCTCTCAACAGGATTTGGTCCCACTGGAGT 366
Db |||
QY 369 CTTGGTGACAAATGGGGATTTATGATCAAGAACAGATTTGCTGATAACCAATACCACTGTT 428
Db |||
QY 367 CTTGGTGACAAATGGGGATTTATGATCAAGAACAGATTTGCTGATAACCAATACCACTGTT 426
Db |||
QY 429 GGTAAATAGGAGCTAAATCTGACAGATTCATGAAACAAAGCGCCATGAAGTTTAACTAG 488
Db |||
QY 427 GGTAAATAGGAGCTAAATCTGACAGATTCATGAAACAAAGCGCCATGAAGTTTAACTAG 486
Db |||
QY 489 GACTGCTCTTCCTGGCTGAGGATTTCAATCCAGAGAGAAATTAATTTGGAGTGCACAAATCC 548
Db |||
QY 487 GACTGCTCTTCCTGGCTGAGGATTTCAATCCAGAGAGAAATTAATTTGGAGTGCACAAATCC 546
Db |||
QY 549 ACAGTACTTAGCTGAGGTTCTTCCAAATGCTGTCAAGCTCAGTAGGTTTTTGTATAAGT 608
Db |||
QY 547 ACAGTACTTAGCTGAGGTTCTTCCAAATGCTGTCAAGCTCAGTAGGTTTTTGTATAAGT 606
Db |||
QY 609 CATAGAGAGAGATCTTTTAAAGAGAGAGTAA--TCAGATTCAGAGCTTTCTCTGATCG 666
Db |||
QY 607 CATAGAGAGAGATCTTTTAAAGAGAGAGTAAATCAGGATTCAGGCTTTCTCTGATCG 666
Db |||
QY 667 AAAAGATTTGGGCGAGAGACATTAAGAGGCTTCATATGACTGA 711
Db |||
QY 667 AAAAGATTTGGGCGAGAGACATTAAGAGGCTTCATATGACTGA 711
Db |||
```

```
RESULT 2
BX099226
LOCUS
DEFINITION
BX099226 Soares fetal liver spleen INFLS.S1 Homo sapiens cDNA clone
IMAGE:9981061023 ; IMAGE:435149, mRNA sequence.
ACCESSION
BX099226
VERSION
BX099226.1 GI:27843984
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 718)
Radelof, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Ebert, L., Schneider, D. and Korn, B.
Human UnigeneSet - RZPD3
Unpublished (2003)
Contact: Ina Rofes
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE:9981061023.
RZPDLiB; I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLiB No.972)
http://www.rzpd.de/CloneCards/cgi-
```

```
bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rofes
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTACACAGGAAACAGCTATGAC.
Location/Qualifiers
1..718
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:9981061023 ; IMAGE:435149"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS.S1"
/notes="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
this is a subtracted version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5',
AACTGGAGAAATAATAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
```

ORIGIN

```
Query Match 96.2%; Score 684; DB 5; Length 718;
Best Local Similarity 100.0%; Pred. No. 1.6e-188;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28 CTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTAGTCCATCTCCTATGCCAAAAT 87
Db |||
QY 1 CTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTAGTCCATCTCCTATGCCAAAAT 60
Db |||
QY 88 CAAAGTCTCGGAAATCCATCATGGAAGTGTGGGCTGCTCAGTGAGTGCAGAGTTCATGAT 147
Db |||
QY 61 CAAAGTCTCGGAAATCCATCATGGAAGTGTGGGCTGCTCAGTGAGTGCAGAGTTCATGAT 120
Db |||
QY 148 TACAAAGAGGAGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 207
Db |||
QY 121 TACAAAGAGGAGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db |||
QY 208 TCTGTGGGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 267
Db |||
QY 181 TCTGTGGGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db |||
QY 268 GGTATTTATTTTCGTACAGACTTAAACAATAGAGAGTCTCCCAAACTTCGCTCGTTGG 327
Db |||
QY 241 GGTATTTATTTTCGTACAGACTTAAACAATAGAGAGTCTCCCAAACTTCGCTCGTTGG 300
Db |||
QY 328 TCATTGGAGAGTCTCAACAGGAGTGTGGTGCACAGTGGAGTCTTGGTGACAAATGGGGAT 387
Db |||
QY 301 TCATTGGAGAGTCTCAACAGGAGTGTGGTGCACAGTGGAGTCTTGGTGACAAATGGGGAT 360
Db |||
QY 388 TATGATCAAGAGACAGTTTGTCTGATACCAAAATACCACTGTTGTAATAGGAGCTAAACTG 447
Db |||
QY 361 TATGATCAAGAGACAGTTTGTCTGATACCAAAATACCACTGTTGTAATAGGAGCTAAACTG 420
Db |||
QY 448 GACCAGATTCATGAAACAAAGGCCCATGAAGTTTAACTAGGACTGCTTTCTGGCTGAG 507
Db |||
QY 421 GACCAGATTCATGAAACAAAGGCCCATGAAGTTTAACTAGGACTGCTTTCTGGCTGAG 480
Db |||
QY 508 GATTTTCAATCCAGAGAGAAATTAATTTGGAGTGCACAAATCCAGGACTTATGCTGCAGGT 567
Db |||
QY 481 GATTTTCAATCCAGAGAGAAATTAATTTGGAGTGCACAAATCCAGGACTTATGCTGCAGGT 540
Db |||
QY 568 TCTTCCAATGCTCAAGCTCAGTAGGTTTTTTTGTAGTGCATAGAGAGAGAGATCTTT 627
Db |||
```

```

Db      541 TCTTCCAACTGTCAGAGCTCAGTAGGTTTTCCTGATGAGTTCATAGAGAGAGATCTTT 600
Qy      628 TTAAGAGAAGGTATCAGATTCAGAGCTTTCCTGATCGGAAAGATTTCGGGCGAGGAACA 687
Db      601 TTAAGAGAAGGTATCAGATTCAGAGCTTTCCTGATCGGAAAGATTTCGGGCGAGGAACA 660
Qy      688 TTAAGAGAGCTTCATATGACTGA 711
Db      661 TTAAGAGAGCTTCATATGACTGA 684

RESULT 3
BG501635 720 bp mRNA linear EST 27-MAR-2001
DEFINITION 602548678F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4654986 5',
mRNA sequence.
ACCESSION BG501635
VERSION BG501635.1 GI:13463152
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 720)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaaps-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1443 row: k column: 19
High quality sequence stop: 719.
Location/Qualifiers
1..720
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4654986"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_61"
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgccctcgcc); Site_2: SfiI (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATCTAGAGCGCGCGGCGGACATG-dT(30)BN-3'
(where B = A, C, G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 95.9%; Score 682.2; DB 4; Length 720;
Best Local Similarity 99.3%; Pred. No. 5.5e-188;
Matches 706; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
Qy 1 ATGCGTCCTCGATCCGGTGAAGGTACTGGTGTGGAGACTCAGGTGTGGAAATCT 60
Db 11 ATGCGTCCTCGATCCGGTGAAGGTACTGGTGTGGAGACTCAGGTGTGGAAATCT 70
Qy 61 TCGTATGCTATCTCTTATGCGCAAAATCAAGTGTGGAAATCCATCATGACTGTGGC 120
Db 71 TCGTATGCTATCTCTTATGCGCAAAATCAAGTGTGGAAATCCATCATGACTGTGGC 130

```

```

Qy      121 TGCTAGTGGATGTCAGAGTTCATGATTACAAAGAGGAAACCCAGAGGAAAGACCTTAC 180
Db      131 TGCTAGTGGATGTCAGAGTTCATGATTACAAAGAGGAAACCCAGAGGAAAGACCTTAC 190
Qy      181 TACATAGAAATATGGGATGTTGGAGGCTCTGTGGGAGTCCAGCAGCGTGAAAGACACA 240
Db      191 TACATAGAAATATGGGATGTTGGAGGCTCTGTGGGAGTCCAGCAGCGTGAAAGACACA 250
Qy      241 AGAGCAGTATTTACAACTCCGTAATGTTATTTTCGTACACGACTTAAACAAATAAG 300
Db      251 AGAGCAGTATTTACAACTCCGTAATGTTATTTTCGTACACGACTTAAACAAATAAG 310
Qy      301 AGTCTCTCCAAACTTTCGTCGTTGTCATTTGAAAGCTCTCAACAGGATTTGGTGCCA 360
Db      311 AGTCTCTCCAAACTTTCGTCGTTGTCATTTGAAAGCTCTCAACAGGATTTGGTGCCA 370
Qy      361 ACTGGAGTCTTGTGACAAATGGGATTTATGATCAAGAACAGTTTGTGTAAACCAATA 420
Db      371 ACTGGAGTCTTGTGACAAATGGGATTTATGATCAAGAACAGTTTGTGTAAACCAATA 430
Qy      421 CCACCTGTTGTAATAGGAGTAAACTGACACAGATTTCATGAAACAAAGCGCCATGAAGTT 480
Db      431 CCACCTGTTGTAATAGGAGTAAACTGACACAGATTTCATGAAACAAAGCGCCATGAAGTT 490
Qy      481 TTAACTAGGACTGCTTCTTCCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 540
Db      491 TTAACTAGGACTGCTTCTTCCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 550
Qy      541 ACAATCCAGGTTACTAGCTGCGAGGTTCTTCCAACTGCTGCAAGCTCAGTAGGTTTTTT 600
Db      551 ACAATCCAGGTTACTAGCTGCGAGGTTCTTCCAACTGCTGCAAGCTCAGTAGGTTTTTT 609
Qy      601 GATAAGCTCATAGAGAGAGATCTTTTAAAGAA--GGTAATCAGATTCCAGGCTTTCC 659
Db      610 GATAAGCTCATAGAGAGAGATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTCC 669
Qy      660 TGATCGGAAAGATTTCGGGCGAGGAAACATTAAAGAGCTTCATTTATGACTG 710
Db      670 TGATCGGAGACAGATTTCGGGCGAGGAACTTAAAGAGCTTCATTTATGACTG 720

RESULT 4
BI333788 732 bp mRNA linear EST 30-JUL-2001
LOCUS 602999279F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5141563 5',
mRNA sequence.
ACCESSION BI333788
VERSION BI333788.1 GI:15018445
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 732)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaaps-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1348 row: m column: 20
High quality sequence stop: 732.
Location/Qualifiers
1..732
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source

```

```
/db xref="taxon:9606"
/clone="TWAGE:5141563"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH MGC 12"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
```

ORIGIN

Query Match	95.8%;	Score 681.4;	DB 4;	Length 732;
Best Local Similarity	99.6%;	Pred. No. 9.4e-188;		
Matches 704;	Conservative 0;	Mismatches 1;	Indels 2;	Gaps 2;

QY 6 GTCCCTGGATCGGGTGAAGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTT 65

Db 1 GTCCCTGGATCGGGTGAAGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTT 60

QY 66 AGTCCATCTCCTATGCCAAATCAAGTCTGGGAAATCCATCATGACTGTGGGCTGCTC 125

Db 61 AGTCCATCTCCTATGCCAAATCAAGTCTGGGAAATCCATCATGACTGTGGGCTGCTC 120

QY 126 AGTGGATGTCAGAGTTCATGATTACAAAGAGGAAACCCAGAGAGAGACCTACTACAT 185

Db 121 AGTGGATGTCAGAGTTCATGATTACAAAGAGGAAACCCAGAGAGAGACCTACTACAT 180

QY 186 AGAATTATGGATGTGGAGGCTCTGTGGGAGTGCAGCAGCGTGAAAGGCAAGAGC 245

Db 181 AGAATTATGGATGTGGAGGCTCTGTGGGAGTGCAGCAGCGTGAAAGGCAAGAGC 240

QY 246 AGTATTCTCAACTCCGTAAATGGTATTATTTTCGTACAGACTTAACAATAGAGTC 305

Db 241 AGTATTCTCAACTCCGTAAATGGTATTATTTTCGTACAGACTTAACAATAGAGTC 300

QY 306 CTCCCAAACTTCGCTGCTTGGTTCATTGGAAGTCTCAACAGGAGTTTGGTGCCAACTGG 365

Db 301 CTCCCAAACTTCGCTGCTTGGTTCATTGGAAGTCTCAACAGGAGTTTGGTGCCAACTGG 360

QY 366 AGTCTTGGTGACAAATGGGATTTATGATCAAGAACAGTGTTCGTGATAACCAATACCCT 425

Db 361 AGTCTTGGTGACAAATGGGATTTATGATCAAGAACAGTGTTCGTGATAACCAATACCCT 420

QY 426 GTTGGTAATAGGACTTAACCTGACAGGATTCATGAAACAAAGCGCCATGAAGTTTAAAC 485

Db 421 GTTGGTAATAGGACTTAACCTGACAGGATTCATGAAACAAAGCGCCATGAAGTTTAAAC 480

QY 486 TAGGACTGCTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGCACAA 545

Db 481 TAGGACTGCTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGCACAA 540

QY 546 TCCACGGTACTTAGCTGACAGGTTCTTCCATGCTGTCAAGCTCAGTAGGTTTTTGATAA 605

Db 541 TCCACGGTACTTAGCTGACAGGTTCTTCCATGCTGTCAAGCTCAGTAGGTTTTTGATAA 600

QY 606 GGTATAGAGAGAGATCTTTTTT-AAGAGAGGTAATCAGATTCAGGCTTCCGTGATC 664

Db 601 GGTATAGAGAGAGATCTTTTTTAAAGAGAGGTAATCAGATTCAGGCTTCCGTGATC 660

QY 665 GGAAGAGATTTGGGGCAGGAGCAATTAAGAGCGCTTCATTATGACTGA 711

Db 661 GGAAGAGATTTGGGGCAGGAGCAATTAAGAGCGCTTCATTATGACTGA 706

RESULT 5

CN361072

LOCUS

DEFINITION

17000600016086 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.

ACCESSION

CN361072

VERSION

CN361072.1 GI:47361006

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 699)

Brandenberger R., Wei, H., Zhang, S., Lei, S., Muraige, J., Fiek, G. J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowski, J. and Stanton, L. W.

Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation

Nat. Biotechnol. 22 (6), 707-716 (2004)

Contact: Brandenberger R

Regenerative Medicine

Geron Corporation

230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658

Fax: 650 473 7760

Email: rbrandenberger@geron.com

Insert Length: 699 Std Error: 0.00.

Location/Qualifiers

1..699

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="embryonic stem cells, DMSO-treated H9 cell line"

/clone_lib="GRN_PREHEP"

/note="Oligo dt primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"

ORIGIN

Query Match	95.7%;	Score 680.6;	DB 7;	Length 699;
Best Local Similarity	99.4%;	Pred. No. 1.6e-187;		
Matches 683;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

QY 1 ATGGCCCTCCCTCGATCGGTGAAGTACTGCTGTGGGAGACTCAGGTGTGGGAAATCT 60

Db 13 ATGGCCCTCCCTCGATCGGTGAAGTACTGCTGTGGGAGACTCAGGTGTGGGAAATCT 72

QY 61 TGTGTAGTCCATCTCCTATGCCAAATCAAGTGTCTGGGAAATCCATCATGAGCTGTGGGC 120

Db 73 TGTGTAGTCCATCTCCTATGCCAAATCAAGTGTCTGGGAAATCCATCATGAGCTGTGGGC 132

QY 121 TGTCTCAGTGGATGTCAGGTTTCATGATTACAAAGAGAGAACCCAGAGAGAGAGACTAC 180

Db 133 TGTCTCAGTGGATGTCAGGTTTCATGATTACAAAGAGAGAACCCAGAGAGAGAGACTAC 192

QY 181 TACATAGAAATTTATGGGATGTTGGAGGCTCTGTGGGAGTCCAGCAGCGTGAAGACACA 240

Db 193 TACATAGAAATTTATGGGATGTTGGAGGCTCTGTGGGAGTCCAGCAGCGTGAAGACACA 252

QY 241 AGAGCAGTATTTCTA CAACTCCGTAATGGTATTTATTTTCGTACACGACTTAAACAATAAG 300

Db 253 AGAGCAGTATTTCTA CAACTCCGTAATGGTATTTATTTTCGTACACGACTTAAACAATAAG 312

QY 301 AAGTCTCTCCCAAACTTGGTGTTCATTTGGGAGCTCTCAACAGGAGTTTGGTGCCA 360

Db 313 AAGTCTCTCCCAAACTTGGTGTTCATTTGGGAGCTCTCAACAGGAGTTTGGTGCCA 372

QY 361 ACTGGAGTCTTGGTGACAAATGGGATTTATGATCAAGAACAGTTCGTGATAACCAATA 420

Db 373 ACTGGAGTCTTGGTGACAAATGGGATTTATGATCAAGAACAGTTCGTGATAACCAATA 432

QY 421 CCACTGTTGGTAATAGGAGCTAAACTGGACCCAGATTTCATGAACAAGCCCATGAAGTT 480

Db 433 CCACTGTTGGTAATAGGAGCTAAACTGGACCCAGATTTCATGAACAAGCCCATGAAGTT 492

QY 481 TTAACCTAGGACTGCTTTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 540

Db 493 TTAACCTAGGACTGCTTTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 552

QY 541 ACAAAATCCACGGTACTTAGCTGCAGGTTCTTCCAAATGCTGTCAAGCTCAGTAGGTTTTT 600

Db 553 ACAAAATCCACGGTACTTAGCTGCAGGTTCTTCCAAATGCTGTCAAGCTCAGTAGGTTTTT 612

```
QY 601 GATAAGGTCATAGAGAAGAGATCTTTTAAAGAGAAGGTAATCAGATTCCAGGCTTTCCT 660
Db 613 GATAAGGTCATAGAGAAGAGATCTTTTAAAGAGAAGGTAATCAGATTCCAGGCTTTCCT 672
QY 661 GATCGGAAAGATTGCGGCGAGGAACA 687
Db 673 GATCGGAAAGATTGCGGCGAGGAACA 699

RESULT 6
CN361073 720 bp mRNA linear EST 16-MAY-2004
LOCUS 328775509 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN361073.1 GI:47361007
ACCESSION EST.
VERSION Homo sapiens (human)
KEYWORDS Homo sapiens
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 720)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 720 Std Error: 0.00.

FEATURES
source
1. .720
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone_lib="GRN_ES"
/notes="oligo dt primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"

ORIGIN
Query Match 95.6%; Score 679.4; DB 7; Length 720;
Best Local Similarity 99.9%; Pred. No. 3.6e-187;
Matches 680; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 31 GTGTGGGAGACTCAGTGTGGGAATCTTCGTAGTCATCTCTATGCCAAATCAA 90
Db 1 GGGTTGGGAGACTCAGTGTGGGAATCTTCGTAGTCATCTCTATGCCAAATCAA 60

QY 91 GTCTGGGAATCCATCATGACTGTGGGCTGCTCAGTGGATGTCAGATTTCATGATTAC 150
Db 61 GTCTGGGAATCCATCATGACTGTGGGCTGCTCAGTGGATGTCAGATTTCATGATTAC 120

QY 151 AAAGAAGGAACCCCGAAGAAGAACCTACTACATAGAAATTTATGGGATGTTGGAGGCTCT 210
Db 121 AAAGAAGGAACCCCGAAGAAGAACCTACTACATAGAAATTTATGGGATGTTGGAGGCTCT 180

QY 211 GTGGGAGTCGACGACGGTGAAAGACACAGAGCAGTATTTACAACTCCGTAATGGT 270
Db 181 GTGGGAGTCGACGACGGTGAAAGACACAGAGCAGTATTTACAACTCCGTAATGGT 240

QY 271 ATTATTTTCGTACACGACTTAACAAATAAGAGTCCTCCCAAACTTGGCTGTGCTCA 330
Db 241 ATTATTTTCGTACACGACTTAACAAATAAGAGTCCTCCCAAACTTGGCTGTGCTCA 300
```

```
QY 331 TTGGAAGCTCTCAACAGGAGATTGGTCCCACTGGAGTCTTGGTGACAAATGGGATTAT 390
Db 301 TTGGAAGCTCTCAACAGGAGATTGGTCCCACTGGAGTCTTGGTGACAAATGGGATTAT 360
QY 391 GATCAAGAACAGTTTGTCTGATAACCAATACCCTGTTGGTAATAGGAGCTAAACTGGAC 450
Db 361 GATCAAGAACAGTTTGTCTGATAACCAATACCCTGTTGGTAATAGGAGCTAAACTGGAC 420
QY 451 CAGATTTCATGAAAACAAAGCGCCATGAAGTTTAACTAGGACTGCTTCTCGGCTGAGGAT 510
Db 421 CAGATTTCATGAAAACAAAGCGCCATGAAGTTTAACTAGGACTGCTTCTCGGCTGAGGAT 480
QY 511 TTCAATCCAGAGAATAATTTGGACTGCAACAATCCACGGTACTTGTAGCTGCAAGTTCT 570
Db 481 TTCAATCCAGAGAATAATTTGGACTGCAACAATCCACGGTACTTGTAGCTGCAAGTTCT 540
QY 571 TCCAATGCTCTCAAGCTCAGTAGCTGTTTGAAGTTCATAGAGAGAGATATCTTTTA 630
Db 541 TCCAATGCTCTCAAGCTCAGTAGCTGTTTGAAGTTCATAGAGAGAGATATCTTTTA 600
QY 631 AGAGAAGGTAATCAGATTCCAGGCTTTCCTGATCGGAAAGATTGGGCGAGGAACATTA 690
Db 601 AGAGAAGGTAATCAGATTCCAGGCTTTCCTGATCGGAAAGATTGGGCGAGGAACATTA 660
QY 691 AAGAGCCTTCATTATGACTCA 711
Db 661 AAGAGCCTTCATTATGACTCA 681

RESULT 7
BQ215951 815 bp mRNA linear EST 02-MAY-2002
LOCUS AGENCOURT_7558711 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6045318
DEFINITION 5', mRNA sequence.
ACCESSION BQ215951
VERSION BQ215951.1 GI:20397351
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 815)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M13289 row: f column: 07
High quality sequence stop: 614.

FEATURES
source
1. .815
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6045318"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 94.9%; Score 674.8; DB 5; Length 815;
```

Best Local Similarity 99.6%; Pred. No. 8.2e-186;
Matches 676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 33 GTTGGAGACTCAGGTGTTGGGAAATCTTCGTTAGTCCATCTCCTATGCGCAAAATCAAGT 92
Db 1 GTTGGAGACTCAGGTGTTGGGAAATCTTCGTTAGTCCATCTCCTATGCGCAAAATCAAGT 60

QY 93 GCTGGAAATCCATCATGAGTGTGGGCTGCTCAGTGGATGTCAGAGTTCATGATTACAA 152
Db 61 GCTGGAAATCCATCATGAGTGTGGGCTGCTCAGTGGATGTCAGAGTTCATGATTACAA 120

QY 153 AGAAGGAACCCAGAGAGAGACCTACTACATAGAAATTTATGGATGTTGGAGGCTCTGT 212
Db 121 AGAAGGAACCCAGAGAGAGACCTACTACATAGAAATTTATGGATGTTGGAGGCTCTGT 180

QY 213 GGGCAGTGCAGCAGCGTGAAGAGCAAGAGCAGTATTCTACAACTCCGTAAATGGTAT 272
Db 181 GGGCAGTGCAGCAGCGTGAAGAGCAAGAGCAGTATTCTACAACTCCGTAAATGGTAT 240

QY 273 TATTTTCGTACAGACTTAACAAATAGAGTCTCCCAAACTTGGGCTGTGTCATT 332
Db 241 TATTTTCGTACAGACTTAACAAATAGAGTCTCCCAAACTTGGGCTGTGTCATT 300

QY 333 GGAAGCTCTCAACAGGGAATTTGGTGCAACTGGAGTCTTGGTGACAAATGGGGATTATGA 392
Db 301 GGAAGCTCTCAACAGGGAATTTGGTGCAACTGGAGTCTTGGTGACAAATGGGGATTATGA 360

QY 393 TCAAGAACAGTTTGTGATAACCAAAATACACCTGTTGGTAATAGGACATAAATCGACCA 452
Db 361 TCAAGAACAGTTTGTGATAACCAAAATACACCTGTTGGTAATAGGACATAAATCGACCA 420

QY 453 GATTCATGAACAAAGGCGCAGTATTTAACTAGGACTGCTTCTGGCTGAGGATTT 512
Db 421 GATTCATGAACAAAGGCGCAGTATTTAACTAGGACTGCTTCTGGCTGAGGATTT 480

QY 513 CAATCAGAGAAATTAATTTGAGCTGCAAAATCCACGGTACTTAGCTGCAGGTTCTTC 572
Db 481 CAATCAGAGAAATTAATTTGAGCTGCAAAATCCACGGTACTTAGCTGCAGGTTCTTC 540

QY 573 CAATGCTGCAAGCTCAGTAGGTTTTTTGATAAGGTCATAGAGAGAGATCTTTTTAAG 632
Db 541 CAATGCTGCAAGCTCAGTAGGTTTTTTGATAAGGTCATAGAGAGAGATCTTTTTAAG 600

QY 633 AGAAGGTAATCAGATTCAGGCTTCTGTGTCGAAAGATTTGGGCGAGGACATTA 692
Db 601 AGAAGGTAATCAGATTCAGGCTTCTGTGTCGAAAGATTTGGGCGAGGACATTA 660

QY 693 GAGCCTTCATTATGACTGA 711
Db 661 GAGCCTTCATTATGACTGA 679

RESULT 8
CR619977
LOCUS CR619977 1867 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DA005YF10 of Neuroblastoma of Homo sapiens (human).
ACCESSION CR619977
VERSION CR619977.1 GI:50500784
KEYWORDS HTC; CDSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1867)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1867)
REFERENCES
AUTHORS Genoscope.

TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : segrif@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1867
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA005YF10"
/tissue_type="Neuroblastoma"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 94.0%; Score 668; DB 3; Length 1867;
Best Local Similarity 100.0%; Pred. No. 1e-183;
Matches 668; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 CAGGTGTTGGGAAATCTTCGTTAGTCCATCTCCTATGCCAAATCAAGTCTGGGAAATC 103
Db 1 CAGGTGTTGGGAAATCTTCGTTAGTCCATCTCCTATGCCAAATCAAGTCTGGGAAATC 60

QY 104 CATCATGAGTCTGTGGGCTCTCAGTGGATGTGAGTTCATGATTAACAAGAGAACCC 163
Db 61 CATCATGAGTCTGTGGGCTCTCAGTGGATGTGAGTTCATGATTAACAAGAGAACCC 120

QY 164 CAGAGAGAGACCTACTACATAGATTTATGGATGTTGGAGCTCTGGGCGAGTGCCA 223
Db 121 CAGAGAGAGACCTACTACATAGATTTATGGATGTTGGAGCTCTGGGCGAGTGCCA 180

QY 224 GCAGGCTGAAAAAGCAAGAGCAGTATTTCTACAACTCCGTAATGTTATTTTCGTAC 283
Db 181 GCAGGCTGAAAAAGCAAGAGCAGTATTTCTACAACTCCGTAATGTTATTTTCGTAC 240

QY 284 AGCATTAAACAATAAGAGTCTCTCCAAAACCTTGGCTGTTGGTCAATGGAGAGCTCTCA 343
Db 241 AGCATTAAACAATAAGAGTCTCTCCAAAACCTTGGCTGTTGGTCAATGGAGAGCTCTCA 300

QY 344 ACAGGGATTTGGTCCCACTGGTGTGACAAATGGGATTTATGATCAAGAACAGT 403
Db 301 ACAGGGATTTGGTCCCACTGGTGTGACAAATGGGATTTATGATCAAGAACAGT 360

QY 404 TTGCTGATAACCAAAATACCACTGTTGTTGTAATAGGAGTAACTGGACCATGATAA 463
Db 361 TTGCTGATAACCAAAATACCACTGTTGTTGTAATAGGAGTAACTGGACCATGATAA 420

QY 464 CAAAGCGCATGAAGTTTAACTAGGACTGCTTTCTGGCTGAGGATTTCAATCCAGAAG 523
Db 421 CAAAGCGCATGAAGTTTAACTAGGACTGCTTTCTGGCTGAGGATTTCAATCCAGAAG 480

QY 524 AATTAATTTGGAGTGCACAAATCCACGGTACTTAGCTCAGGTTCTTCCAAATGCTGCA 583
Db 481 AATTAATTTGGAGTGCACAAATCCACGGTACTTAGCTCAGGTTCTTCCAAATGCTGCA 540

QY 584 AGCTCAGTAGGTTTTTTGATAAGGTCATAGAGAAGAGATACTTTTTAAGAGAGGTAATC 643
Db 541 AGCTCAGTAGGTTTTTTGATAAGGTCATAGAGAAGAGATACTTTTTAAGAGAGGTAATC 600

QY 644 AGATTCAGGCTTTCTGATCGGAAAGAATTTGGGCGAGGAAACATTTAAAGAGGCTTCATT 703
Db 601 AGATTCAGGCTTTCTGATCGGAAAGAATTTGGGCGAGGAAACATTTAAAGAGGCTTCATT 660

QY 704 ATGACTGA 711
Db 661 ATGACTGA 668

RESULT 9

```
AL516309
LOCUS AL516309 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone linear EST 23-MAR-2004
DEFINITION CS0DA005YF10 5-PRIME, mRNA sequence.
ACCESSION AL516309
VERSION AL516309.3 GI:45652953
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 797)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:30490812.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1912.r,
and it belongs to a clone representative of this cluster. For more
information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/cdna?s=CS0DA005YF10&c=1912.r.
FEATURES
source
1..797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA005YF10"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 93.3%; Score 663.6; DB 1; Length 797;
Best Local Similarity 98.8%; Pred. No. 1.5e-182;
Matches 660; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
QY 44 CAGGTGTGGGAATCTTCGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAATC 103
DB 1 CAGGTGTGGGAATCTTCGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAATC 60
QY 104 CATCATGGACTGTGGGCTGCTCAGTGGATGTCAGATTCATGATTACAAAGAGGACCC 163
DB 61 CATCATGGACTGTGGGCTGCTCAGTGGATGTCAGATTCATGATTACAAAGAGGACCC 120
QY 164 CAGAAGAGAGACCTACTACATAGAATTATGGGATGTGGAGGCTCTGTGGGAGTCCCA 223
DB 121 CAGAAGAGAGACCTACTACATAGAATTATGGGATGTGGAGGCTCTGTGGGAGTCCCA 180
QY 224 GCAGCGTGAAGACCAAGAGCAGTATCTTACAACTCCGTAATGGTATATTTCGTAC 283
DB 181 GCAGCGTGAAGACCAAGAGCAGTATCTTACAACTCCGTAATGGTATATTTCGTAC 240
QY 284 ACCACTTAACAAATAGAACTCTCCCAAACTTGGTGGATTCATTCGAGCTCTCA 343
DB 241 ACCACTTAACAAATAGAACTCTCCCAAACTTGGTGGATTCATTCGAGCTCTCA 300
QY 344 ACAGGGATTTGGTCCCACTGGAGTCTTGGTGACAAATGGGATTTATGATCAAGAACGT 403
DB 301 ACAGGGATTTGGTCCCACTGGAGTCTTGGTGACAAATGGGATTTATGATCAAGAACGT 360
QY 404 TTGCTGATAACCAATACCACTGTGTTGTAATAGGGACTAACTGGACCAAGATTATGAAA 463
```

```
Db 361 TTGCTGATAACCAATACCACTGTGTTGTAATAGGGACTAACTGGACCAAGATTATGAAA 420
QY 464 CAAAGCGCCATGAAGTTTAACTAGAGTCTGTTTCCCTGGCTGAGGATTTCAATCCAGAAG 523
DB 421 CAAAGCGCCATGAAGTTTAACTAGAGTCTGTTTCCCTGGCTGAGGATTTCAATCCAGAAG 480
QY 524 AAATTAATTTGGACTGACCAAAATCCACGGTACTAGCTGCAGGTCTTCCATGCTGTCA 583
DB 481 AAATTAATTTGGACTGACCAAAATCCACGGTACTAGCTGCAGGTCTTCCATGCTGTCA 540
QY 584 AGCTCAGTAGTGTGTTTGAATAAGTCTATAGAGAAGAGATCTTTTAAAGAGAGGTAAATC 643
DB 541 AGCTCAGTAGTGTGTTTGAATAAGTCTATAGAGAAGAGATCTTTTAAAGAGAGGTAAATC 600
QY 644 AGATTCAGGCTTTTCCCTGATCGGAAAGATTTGGGCGAGGAACATTAAAGAGCCTTCATT 703
DB 601 AGATTCAGGCTTTTCCCTGATCGGAAAGATTTGGGCGAGGAACATTAAAGAGCCTTCATT 660
QY 704 ATGACTGA 711
DB 661 ATGACTGA 668
RESULT 10
BM451532
LOCUS BM451532 989 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT 6395013 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493017
5', mRNA sequence.
ACCESSION BM451532
KEYWORDS EST.
SOURCE BM451532.1 GI:18500572
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 989)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12116 row: a column: 18
High quality sequence stop: 652.
FEATURES
source
1..989
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5493017"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
ORIGIN
Query Match 92.0%; Score 654.4; DB 4; Length 989;
Best Local Similarity 99.7%; Pred. No. 8e-180;
Matches 666; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 45 AGGTGTGGGAATCTTCGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAATCC 104
DB 182 AGGTGTGGGAATCTTCGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAATCC 241
```

```
QY 105 ATCATGGAGCTGGGCTGCTCAGTGATGTCAGAGTTCATGATTAACAAGAGGACCCC 164
Db 242 ATCATGGAGCTGGGCTGCTCAGTGATGTCAGAGTTCATGATTAACAAGAGGACCCC 301
QY 165 AGAAGAGAGACCTACTACATAGAAATATGGGATGTTGGAGGCTCTGTGGCAGTGCCAG 224
Db 302 ABAAGAGAGACCTACTACATAGAAATATGGGATGTTGGAGGCTCTGTGGCAGTGCCAG 361
QY 225 CAGCGTGAAAAGCACAAGAGCAGTATTCACAACTCCGTAATATGGTATTAATTTTCGTACA 284
Db 362 CAGCGTGAAAAGCACAAGAGCAGTATTCACAACTCCGTAATATGGTATTAATTTTCGTACA 421
QY 285 CGACTTAAACAATAAGAGCTCTCCCAAACTGGCTGCTGGTCTATTTGGAGCTCTCAA 344
Db 422 CGACTTAAACAATAAGAGCTCTCCCAAACTGGCTGCTGGTCTATTTGGAGCTCTCAA 481
QY 345 CAGGGAATTTGGTGCCAACTGGAGTCTTTGGTGACAAATGGGGATTAATGATCAAGAACAGTT 404
Db 482 CAGGGAATTTGGTGCCAACTGGAGTCTTTGGTGACAAATGGGGATTAATGATCAAGAACAGTT 541
QY 405 TGGTGATAACCAATACCACTGTTGGTAATAGGACTTAACTGGACAGATTCATGAAC 464
Db 542 TGGTGATAACCAATACCACTGTTGGTAATAGGACTTAACTGGACAGATTCATGAAC 601
QY 465 AAAGCCCATGAATTTAACTAGGAGTCTTTCTGGCTGAGGATTTCAATCCAGAAGA 524
Db 602 AAAGCCCATGAATTTAACTAGGAGTCTTTCTGGCTGAGGATTTCAATCCAGAAGA 661
QY 525 AATTAATTTGGAGTGCACAAATCCAGTACTTTAGCTGCAGGTTCTTCAATGCTGCA 584
Db 662 AATTAATTTGGAGTGCACAAATCCAGTACTTTAGCTGCAGGTTCTTCAATGCTGCA 721
QY 585 GCTCAGTAGGTTTTTGATAAGGTCATAGAGAGAGATCTTTTAAAGAGAGGTAATCA 644
Db 722 GCTCAGTAGGTTTTTGATAAGGTCATAGAGAGAGATCTTTTAAAGAGAGGTAATCA 781
QY 645 GATTCAGGC-TTCTCTGATCGGAAAGATTTGGGCGAGGAACATTAAGAGGCTTCATT 703
Db 782 GATTCAGGCTTTTCTGTGATCGGGAAGATTTGGGCGAGGAACATTAAGAGGCTTCATT 841
QY 704 ATGACTGA 711
Db 842 ATGACTGA 849

RESULT 11
BG258348
LOCUS
DEFINITION
602379905F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4510560 5',
mRNA sequence.
ACCESSION
BG258348
VERSION
BG258348.1 GI:12768164
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 902)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10392 row: j column: 01
High quality sequence stop: 648.
```

```
FEATURES
source
Location/Qualifiers
1..902
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4510560"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 91.4%; Score 650.2; DB 4; Length 902;
Best Local Similarity 98.7%; Pred. No. 1.3e-178;
Matches 666; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 19 GTGAAGGTAAGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTAGTCCATCTCCTA 78
Db 1 GTGAAGGTAAGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTAGTCCATCTCCTA 60
QY 79 TCCAAAATCAAGTGTGGGAAATCCATCATGTGGCTGTGGCTGTCTCAGTGAATGTCAGA 138
Db 61 TCCAAAATCAAGTGTGGGAAATCCATCATGTGGCTGTGGCTGTCTCAGTGAATGTCAGA 120
QY 139 GTTCATGATTACAAAGAGAAACCCAGAGAGAGAACCTTACTACATAGAAATTTATGGAT 198
Db 121 GTTCATGATTACAAAGAGAAACCCAGAGAGAGAACCTTACTACATAGAAATTTATGGAT 180
QY 199 GTTGAGGCTCTGTGGGAGTGCAGCGGTGAAAAGCACAAGACAGATTTCTCAAC 258
Db 181 GTTGAGGCTCTGTGGGAGTGCAGCGGTGAAAAGCACAAGACAGATTTCTCAAC 240
QY 259 TCCGTAATGTGATTTATTTTCGTACACGACTTAAACAAATAGAGTCTCCCAAACTTG 318
Db 241 TCCGTAATGTGATTTATTTTCGTACACGACTTAAACAAATAGAGTCTCCCAAACTTG 300
QY 319 CGTGTGTGTCATTTGGAAGCTCTCAACAGGAGTTTGGTGCCAACTGGAGTCTTGGTGACA 378
Db 301 CGTGTGTGTCATTTGGAAGCTCTCAACAGGAGTTTGGTGCCAACTGGAGTCTTGGTGACA 360
QY 379 AATGGGATATGATCAAGAACAGTTTGTCTGATAACCAATACCAGTGTGTAATAGG 438
Db 361 AATGGGATATGATCAAGAACAGTTTGTCTGATAACCAATACCAGTGTGTAATAGG 420
QY 439 ACTAACTGGACAGATTCATGAAACAAAGCGCATGAAAGTTTAACTAGGACTGCTTTC 498
Db 421 ACTAACTGGACAGATTCATGAAACAAAGCGCATGAAAGTTTAACTAGGACTGCTTTC 480
QY 499 CTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGAGTGCACAAATCCACGGTACTTA 558
Db 481 CTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGAGTGCACAAATCCACGGTACTTA 540
QY 559 GCTGCAAGTCTTCCATGCTGATGCTCAGTGTGTTTTCATAGGTCATAGAGAG 618
Db 541 GCTGCAAGTCTTCCATGCTGATGCTCAGTGTGTTTTCATAGGTCATAGAGAG 600
QY 619 AGATACCTTTTAAAGAGAGGTAATCAGATTCAGAGGCTTTCCTGATCGGAA-AAAGTTTGG 677
Db 601 AGATACCTTTTAAAGAGAGGTAATCAGATTCAGAGGCTTTCCTGATCGGAA-AAAGTTTGG 660
QY 678 GGCAGGAACATTAATA 692
Db 661 GGCAGGAACATTAATA 675

RESULT 12
CK455329
LOCUS
DEFINITION
915906 MARC 4P1G Sus scrofa cDNA 5', mRNA linear EST 12-JAN-2004
```

```
ACCESSION CK455329
VERSION CK455329.1 GI:40802543
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
AUTHORS Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE Porcine EST collection using a normalized library constructed from
        embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
        USDA, ARS, US Meat Animal Research Center
        PO Box 166, Clay Center, NE 68933-0166, USA
        Tel: 402 762 4366
        Fax: 402 762 4390
        Email: smith@email.marc.usda.gov
        Single pass sequencing. Bases called with phred v0.020425.c and
        trimmed with the aid of the trim_ait option. Vector identified with
        cross match v0.990329.
        Plate: TWB8028 row: B column: 12
Seq primer: GTAATACGACTCACTATAGG.
          Location/Qualifiers
            1..855
              /organism="Sus scrofa"
              /mol_type="mRNA"
              /db_xref="taxon:9823"
              /tissue_type="pooled"
              /lab_host="DH10B"
              /clone_lib="MARC 4P1G"
              /notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
              Library made with combined RNA from day-10, day-13,
              day-15, day-25, and day-30 whole embryos."

FEATURES
    source
        1..855
            /organism="Sus scrofa"
            /mol_type="mRNA"
            /db_xref="taxon:9823"
            /tissue_type="pooled"
            /lab_host="DH10B"
            /clone_lib="MARC 4P1G"
            /notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
            Library made with combined RNA from day-10, day-13,
            day-15, day-25, and day-30 whole embryos."

ORIGIN
Query Match          91.4%; Score 650; DB 7; Length 855;
Best Local Similarity 95.0%; Pred. No. 1.5e-178;
Matches 671; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 6 GTCCCTGGATCGGTCAGAGTACTGGTGTGGAGACTCAGGTGTGGGAAATCTTCGTT 65
DB 1 GTCCCTGGATCGGTCAGAGTACTGGTGTGGAGACTCAGGTGTGGGAAATCTTCAC 60

QY 66 AGTCCATCTCTATGCCAAATCAAGTGTGGGAAATCCATCATGCACTGTGGCTGCTC 125
DB 61 AGTTCACTCTTAATGCCAAATCAAGTGTGGGAAATCCATCATGCACTGTGGCTGCTC 120

QY 126 AGTGGATGTCAGAGTTTCATGATTACAAAGAGGAAGAACCCAGAGAGAGACCTACTACAT 185
DB 121 AGTAGATGTCAGAGTTTCATGACTACAAAGAGGAAGAACCCAGAGAGAGACCTACTATAT 180

QY 186 AGAATTATGGGATGTGGAGGCTCTGTGGCAGTGCAGCAGGTCGAAAGACACAGAGC 245
DB 181 AGAATTATGGGATGTGGAGGCTCTGTGGCAGTGCAGCAGTGTGAAAAGACACAGAGC 240

QY 246 AGTATTCTACAACCTCCGTAATGTTATTTTCGTACACGACTTAAACAATAAGAGTC 305
DB 241 AGTATTCTACAACCTCTGTAATGTTATTTTAGTATACGACTTAAACAATAAGAGTC 300

QY 306 CTCCCAAAATCTGGCTGTTGGTCAATTTGGAAGCTCTCAACAGGGAATTTGGTGCACCTGG 365
DB 301 ATCCCAAAATTTGATGTTGGTCAATTTGGAAGCTCTCAACAGGGAATCTGGTGCACCTGG 360

QY 366 AGTCTTGGTCAAAATGGGGAATATGATCAAGAACAGTGTGCTGATACCAATACCACT 425
DB 361 AGTCTTGGTCAAAATGGGGAATATGATCAAGAACAGTGTGCTGATACCAATACCACT 420

QY 426 GTTGGTAATAGGCACTAAACCTGGACAGATTCATGAACAAAGCCCATCAAGTTTAAAC 485
DB 421 GTTGGTAATAGGCACTAAACCTGGATCATGATTCATGAACAAAGCCCATCAAGTTTAAAC 480
```

```
486 TAGGACTGCTTTCCTGCTGAGGATTTCAATCCAGAAGAAATTAATTGGACTGCACAA 545
481 TAGGACTGCTTTCCTGCTGAGGATTTCAATGAGAAGAGATTAATTGGATTGCACAA 540
546 TCCACAGTACTTACGTCAGGTTCTTCCCAATGCTCAAGCTCAGTAGAGTGTTCGATAA 605
541 TCCACAGTACTTACGTCAGGTTCTTCCCAATGCTCAAGCTCAGTAGAGTGTTCGATAA 600
606 GGTCTATAGAGAAGAGATACATTTTAAAGAGAAGGTAATCAGATTCAGAGCTTCCTGATCG 665
601 GGTCTATAGAGAAGAGATACATTTTAAAGAGAAGGTAATCAGATTCAGATTCAGAGCTTCCTGATCG 660
666 GAAAGAGTTTGGGCGAGGACATTAAGAGCCTTCATTATGACTGA 711
661 AAAAAAGTTTGGAGGAGGAACATTAAGAGCTTACATTATGACTGA 706

RESULT 13
BI759699 785 bp mRNA linear EST 25-SEP-2001
LOCUS G03045561F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186006 5',
DEFINITION mRNA sequence.
ACCESSION BI759699
VERSION BI759699.1 GI:15751277
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 785)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
        Email: cgabbs-r@mail.nih.gov
        Tissue Procurement: Life Technologies, Inc.
        cDNA Library Preparation: Life Technologies, Inc.
        cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
        DNA Sequencing by: Incyte Genomics, Inc.
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LLNL at:
        http://image.llnl.gov
        Plate: LLAM11464 row: i column: 15
        High quality sequence start: 13
        High quality sequence stop: 761.

FEATURES
    source
        1..785
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5186006"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_116"
            /notes="Organ: pooled colon, kidney, stomach; Vector:
            pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
            source anonymous pool of 3 colons, age 26 yo male, 49 yo
            female, 71 yo male colon; 46 yo male kidney, and pool of 2
            stomachs, 62 yo male and 70 yo female. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.4 Kb,
            insert size range 1-3 Kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            023. Note: this is a NIH_MGC Library."

ORIGIN
Query Match          90.6%; Score 644; DB 4; Length 785;
Best Local Similarity 97.8%; Pred. No. 8.3e-177;
Matches 696; Conservative 0; Mismatches 10; Indels 6; Gaps 4;

QY 3 GGGCTCCTGGATCGGTCAGGTCAGTGTGGGAGACTCAGGTGTGGGAAATCTT- 61
DB 1 GGGCTCCTGGATCGGTCAGGTCAGTGTGGGAGACTCAGGTGTGGGAAATCTTG 60
```


SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 959)
AUTHORS NTH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10089 Row: m Column: 11
High quality sequence stop: 665.
Location/Qualifiers
1..959
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4394290"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 87"
/notes="Organ: breast; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN

Query Match 88.3%; Score 628; DB 4; Length 959;
Best Local Similarity 99.4%; Pred. No. 4.2e-172;
Matches 662; Conservative 0; Mismatches 0; Indels 4; Gaps 3;
QY 48 TGTGGGAATCTTCGTTAGTCCATCTCTATGCAAAATCAAGTCTGGGAATCCATC 107
DB 1 TGTGGGAATCTTCGTTAGTCCATCTCTATGCAAAATCAAGTCTGGGAATCCATC 60
QY 108 ATGACTGTGGCTGCTCAGTGGATGTCAGATTCATGATTAACAAGAAACCCCGA 167
DB 61 ATGAC - TGGGCTGCTCAGTGGATGTCAGATTCATGATTAACAAGAAACCCCGA 118
QY 168 AGAAGACCTACTACATAGAAATATGGATGTTGGAGGCTCTGTGGGAGTGCACAG 227
DB 119 AGAAGACCTACTACATAGAAATATGGATGTTGGAGGCTCTGTGGGAGTGCACAG 178
QY 228 CGTGAAGCACAAGACAGTATCTCACTCGTAAATGGTATATTTTCGTACACGA 287
DB 179 CGTGAAGCACAAGACAGTATCTCACTCGTAAATGGTATATTTTCGTACACGA 238
QY 288 CTTAACAATAAGAAGTCTCCCAAACTTCGCTGTTGGTCAATGGAAGCTCTCAACAG 347
DB 239 CTTAACAATAAGAAGTCTCCCAAACTTCGCTGTTGGTCAATGGAAGCTCTCAACAG 298
QY 348 GGATTTGGTGCACACTGGAGTCTTGGTGACAAATGGGGATATGATCAAGAACAGTTTGC 407
DB 299 GGATTTGGTGCACACTGGAGTCTTGGTGACAAATGGGGATATGATCAAGAACAGTTTGC 358
QY 408 TGATAACCAATACCTGTTGGTAATAGGACATAAATGGACAGATTCATGAACAAA 467
DB 359 TGATAACCAATACCTGTTGGTAATAGGACATAAATGGACAGATTCATGAACAAA 418
QY 468 GCGCCATCAAGTTTAACTAGGACTGCTTTCTCGGCTGAGGATTTCAATCCAGAGAAAT 527
DB 419 GCGCCATCAAGTTTAACTAGGACTGCTTTCTCGGCTGAGGATTTCAATCCAGAGAAAT 478
QY 528 TAATTTGGACTGCAAAATCCACGGTACTTAGCTGCAGGTTCTTCCAAATGCTGCAAGCT 587

Db 479 TAATTTGGACTGCAAAATCCACGGTACTTAGCTGCAGGTTCTTCCAAATGCTGCAAGCT 538
QY 588 CAGTAGGTTTTTTTGATAAGGTCATAGAGAAGAGATAC-TTTTTAAGAGAAGGTAATCAGA 646
DB 539 CAGTAGGTTTTTTTGATAAGGTCATAGAGAAGAGATAC-TTTTTAAGAGAAGGTAATCAGA 598
QY 647 TTCCAGGCTTTCTCTGATCGG-AAAGAGATTTGGGCGAGAACATTAAGAGGCTTCATAT 705
DB 599 TTCCAGGCTTTCTCTGATCGGAAAAAAGATTTGGGCGAGAACATTAAGAGGCTTCATAT 658
QY 706 GACTGA 711
DB 659 GACTGA 664

Search completed: June 30, 2005, 16:53:34
Job time : 2655.06 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 14:24:03 ; Search time 3257.89 Seconds
(without alignments)
10574.845 Million cell updates/sec

Title: US-09-945-173-3
Perfect score: 711
Sequence: 1 atggcctcctgactgggt.....agagccttcattatgactga 711

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	711	100.0	711	6	AX465671 Sequence
2	711	100.0	1694	6	AX465669 Sequence
3	711	100.0	3362	9	BSX647559 Homo sapi
4	711	100.0	3449	9	BSX647593 Homo sapi
5	709.4	99.8	3352	9	BSX647664 Homo sapi
6	707.8	99.5	1999	9	BC020832 Homo sapi
7	587.8	82.7	2064	10	BC050194 Homo sapi
8	484.2	68.1	3327	5	AJ720020 Gallus ga
9	435.4	61.2	2550	9	AK025772 Homo sapi
10	418.2	58.8	1064	5	BC073035 Xenopus l
11	396.6	55.8	1040	5	BC078191 Danio rer
12	395	55.6	1291	5	AF391440
13	365.8	51.4	385	6	BD076273
14	223	31.4	223	6	CQ719549 Sequence
15	154.2	21.7	65838	9	AC133474 Homo sapi
16	154.2	21.7	146010	9	AC009289 Homo sapi
17	139.2	19.6	66009	2	AC101250 Mus muscu
18	139.2	19.6	155418	2	AC117754 Mus muscu
19	135	19.0	119730	9	AC117472 Homo sapi

20	135	19.0	188485	2	AC027454	AC027454 Homo sapi
21	128.6	18.1	245326	2	AC107288	AC107288 Rattus no
22	128.6	18.1	266597	2	AC119722	AC119722 Rattus no
c 23	102.2	14.4	165836	2	AC138153	AC138153 Danio rer
24	102.2	14.4	174601	5	BX571724	BX571724 Zebrafish
c 25	102.2	14.4	180928	2	AC139727	AC139727 Danio rer
c 26	81.6	11.5	87017	2	AC140043	AC140043 Mus muscu
c 27	77.6	10.9	29949	2	AC112488	AC112488 Homo sapi
c 28	77.6	10.9	179351	2	AC143587	AC143587 Macaca mu
c 29	66.4	9.3	919	6	CQ581943	CQ581943 Sequence
30	66.4	9.3	934	3	AY070945	AY070945 Drosophil
31	66.4	9.3	1087	6	AR508948	AR508948 Sequence
32	56.8	8.0	3048	6	CQ581942	CQ581942 Sequence
33	56.8	8.0	3215	6	CQ591803	CQ591803 Sequence
34	56.8	8.0	3888	6	CQ591806	CQ591806 Sequence
c 35	56.8	8.0	8610	2	AC012956	AC012956 Drosophil
c 36	56.8	8.0	177576	3	AC022349	AC022349 Drosophil
c 37	56.8	8.0	179077	3	AC018489	AC018489 Drosophil
c 38	56.8	8.0	300469	3	AE003503	AE003503 Drosophil
c 39	49.2	6.9	188485	2	AC027454	AC027454 Homo sapi
40	47.6	6.7	138	6	CQ736278	CQ736278 Sequence
41	46	6.5	1541	8	AK068217	AK068217 Oryza sat
42	45.2	6.4	1438	8	AK071529	AK071529 Oryza sat
c 43	45	6.3	65009	2	AC101250	AC101250 Mus muscu
c 44	44.4	6.2	183353	9	AC087071	AC087071 Homo sapi
45	44.4	6.2	185273	9	AC073320	AC073320 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AX465671 711 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0218425.
ACCESSION AX465671
VERSION AX465671.1 GI:21899931
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Meyers, R.A.
TITLE 47324, a human g-protein and uses therefor
JOURNAL Patent: WO 0218425-A 3 07-MAR-2002;
Millennium Pharmaceuticals (US)
FEATURES
source Location/Qualifiers
1..711
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN	Query Match	Best Local Similarity	Matches	Score	DB 6;	Length	711;	DB 6;	Length	711;	Indels	0;	Gaps	0;
Qy	1	ATGCGCTCCCTCGGATCGGTTGAAGTACTGGTGTGGAGACTCAGGTGTTGGGAAATCT	60	100.0%;	100.0%;	711;	DB 6;	100.0%;	100.0%;	711;	0;	0;	0;	0;
Db	1	ATGCGCTCCCTCGGATCGGTTGAAGTACTGGTGTGGAGACTCAGGTGTTGGGAAATCT	60	100.0%;	100.0%;	711;	DB 6;	100.0%;	100.0%;	711;	0;	0;	0;	0;
Qy	61	TCGTTAGTCCATCTCTATGCCAAAATCAAGTGTCTGGGAAATCCATCATGAGCTGTGGGC	120	100.0%;	100.0%;	711;	DB 6;	100.0%;	100.0%;	711;	0;	0;	0;	0;
Db	61	TCGTTAGTCCATCTCTATGCCAAAATCAAGTGTCTGGGAAATCCATCATGAGCTGTGGGC	120	100.0%;	100.0%;	711;	DB 6;	100.0%;	100.0%;	711;	0;	0;	0;	0;
Qy	121	TGCTCAGTGTGATGTGAGTTTCATGATTACAAGAGGAAACCCAGAGAGAACCTTAC	180	100.0%;	100.0%;	711;	DB 6;	100.0%;	100.0%;	711;	0;	0;	0;	0;
Db	121	TGCTCAGTGTGATGTGAGTTTCATGATTACAAGAGGAAACCCAGAGAGAACCTTAC	180	100.0%;	100.0%;	711;	DB 6;	100.0%;	100.0%;	711;	0;	0;	0;	0;
Qy	181	TACATAGAAATATGGGATGTTGGAGCTCTGTGGGAGTCCAGCAGCGTGAAGGACA	240	100.0%;	100.0%;	711;	DB 6;	100.0%;	100.0%;	711;	0;	0;	0;	0;
Db	181	TACATAGAAATATGGGATGTTGGAGCTCTGTGGGAGTCCAGCAGCGTGAAGGACA	240	100.0%;	100.0%;	711;	DB 6;	100.0%;	100.0%;	711;	0;	0;	0;	0;

QY	241	AGAGCAGTATTCTACAACTCCGTAATAATGGTATTAATTTTGGTACACGACTTAACAAATAAG	300
DB	241	AGAGCAGTATTCTACAACTCCGTAATAATGGTATTAATTTTGGTACACGACTTAACAAATAAG	300
QY	301	AAGTCCTCCCAAACTTGGCTGGTCAATCGAAGCTCTCAACAGGGAATTTGGTGCCA	360
DB	301	AAGTCCTCCCAAACTTGGCTGGTCAATCGAAGCTCTCAACAGGGAATTTGGTGCCA	360
QY	361	ACTGGAGTCTTGGTGCAAAATGGGATTAATGATCAAGAACAGTTTCTGATTAACCAATA	420
DB	361	ACTGGAGTCTTGGTGCAAAATGGGATTAATGATCAAGAACAGTTTCTGATTAACCAATA	420
QY	421	CACTGTTGGTAATAGGACTAACTGGACAGATTATGATCAAGAACAGTTTCTGATTAACCAATA	480
DB	421	CACTGTTGGTAATAGGACTAACTGGACAGATTATGATCAAGAACAGTTTCTGATTAACCAATA	480
QY	481	TTAACTAGGACTCTTCTCGCTGAGGATTTCAATCCAGAACAAATTAATTTGGACTGC	540
DB	481	TTAACTAGGACTCTTCTCGCTGAGGATTTCAATCCAGAACAAATTAATTTGGACTGC	540
QY	541	ACAAATCCAGGTAATAGCTGAGGTTCTTCCAAATGCTGCTCAAGCTCAGTAGGTTTTT	600
DB	541	ACAAATCCAGGTAATAGCTGAGGTTCTTCCAAATGCTGCTCAAGCTCAGTAGGTTTTT	600
QY	601	GATAAGTCTATAGAGAGAGATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTCT	660
DB	601	GATAAGTCTATAGAGAGAGATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTCT	660
QY	661	GATCGGAAGAATTTGGGCGAGCAATTAAGAGGCTTCATTTATGACTGA	711
DB	661	GATCGGAAGAATTTGGGCGAGCAATTAAGAGGCTTCATTTATGACTGA	711
RESULT 2			
AX465669			
LOCUS			
Sequence 1 from Patent WO0218425.			
AX465669			
VERSION			
AX465669.1 GI:21899929			
KEYWORDS			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1			
Meyers, R.A.			
47324, a human g-protein and uses therefor			
Patent: WO 0218425-A 1 07-MAR-2002;			
JOURNAL			
Millennium Pharmaceuticals (US)			
FEATURES			
source			
1. .1694			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
19. .729			
/notes="unnamed protein product"			
/codon_start=1			
/protein_id="CAD42379.1"			
/db_xref="GI:21899930"			
/translation="MASLRKVLVLGDSGVKSLVHLLCQNOVLGNPSWTVGCSVD			
VRVHDYKGFPEEKTYIELVDVGGSVGSVKSRAVFNYSVNGIIFVHDLTNKKS			
SQLRWSLEALNRDLPVPTGLVNGDYPQEQPADNQIPLVLGTLKLDGTHETKHEV			
LRTAFLEADFNPEELNLDCTNPRYLAAGSSNAVKLSRFPDKVIEKRYFLREGNQIIPG			
FPDRKRFAGATLKLSDH"			
ORIGIN			
Query Match 100.0%; Score 711; DB 6; Length 1694;			
Best Local Similarity 100.0%; Pred. No. 4.5e-198;			
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGCGCTCCCTGATCGGGTGAAGGTACTCGTGTGGAGACTCAGTGTGGGAAATCT	60

DB	19	ATGCGCTCCCTGGATCGGGTGAAGGTACTCGTGTGGAGACTCAGTGTGGGAAATCT	78
QY	61	TCGTTAGTCCATCTCTTATGCCAAATCAAGTGTGGGAAATCCATCATGGACTGTGGGC	120
DB	79	TCGTTAGTCCATCTCTTATGCCAAATCAAGTGTGGGAAATCCATCATGGACTGTGGGC	138
QY	121	TCGCTCAGTGGATGTCAGAGTTCATGATTACAAGAGGAACCCCAAGAGAGACCTAC	180
DB	139	TCGCTCAGTGGATGTCAGAGTTCATGATTACAAGAGGAACCCCAAGAGAGACCTAC	198
QY	181	TACATAGAAATTAATGGAGGCTCTGTGGGAGTCCAGCAGAGCGTGAAAAGCACA	240
DB	199	TACATAGAAATTAATGGAGGCTCTGTGGGAGTCCAGCAGAGCGTGAAAAGCACA	258
QY	241	AGAGCAGTATTCTCAAACTCCGTAATAATGGTATTAATTTTGGTACACGACTTAACAAATAAG	300
DB	259	AGAGCAGTATTCTCAAACTCCGTAATAATGGTATTAATTTTGGTACACGACTTAACAAATAAG	318
QY	301	AAGTCCTCCCAAACTTGGCTGGTCAATTTGGAGCTCTCAACAGGGAATTTGGTGCCA	360
DB	319	AAGTCCTCCCAAACTTGGCTGGTCAATTTGGAGGCTCTCAACAGGGAATTTGGTGCCA	378
QY	361	ACTGAGTCTTGGTGACAAATGGGATTAATGATCAAGAACAGTTTCTGATTAACCAATA	420
DB	379	ACTGAGTCTTGGTGACAAATGGGATTAATGATCAAGAACAGTTTCTGATTAACCAATA	438
QY	421	CACTGTTGGTAAATAGGACTAACTGGACAGATTATGATCAAGAACAGTTTCTGATTAACCAATA	480
DB	439	CACTGTTGGTAAATAGGACTAACTGGACAGATTATGATCAAGAACAGTTTCTGATTAACCAATA	498
QY	481	TTAACTAGGACTCTTCTCGCTGAGGATTTCAATCCAGAACAAATTAATTTGGACTGC	540
DB	499	TTAACTAGGACTCTTCTCGCTGAGGATTTCAATCCAGAACAAATTAATTTGGACTGC	558
QY	541	ACAAATCCAGGTAATAGCTGAGGTTCTTCCAAATGCTGCTCAAGCTCAGTAGGTTTTT	600
DB	559	ACAAATCCAGGTAATAGCTGAGGTTCTTCCAAATGCTGCTCAAGCTCAGTAGGTTTTT	618
QY	601	GATAAGTCTATAGAGAGAGATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTCT	660
DB	619	GATAAGTCTATAGAGAGAGATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTCT	678
QY	661	GATCGGAAGAATTTGGGCGAGCAATTAAGAGGCTTCATTTATGACTGA	711
DB	679	GATCGGAAGAATTTGGGCGAGCAATTAAGAGGCTTCATTTATGACTGA	729
RESULT 3			
HSM807705			
LOCUS			
Homo sapiens mRNA; cDNA DKFp313A1611 (from clone DKFp313A1611).			
BX647559			
VERSION			
BX647559.1 GI:34366716			
KEYWORDS			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 3362)			
AUTHORS			
Koehler, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,			
Fobo, G., Han, M. and Wiemann, S.			
The German Human cDNA Consortium			
Direct Submission			
Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764			
Neuhberg, GERMANY			
CLONE			
Neuhberg, GERMANY			
Research Center (DKPZ); Email s.wiemann@kfz-heidelberg.de;			
sequenced by BWFZ (Biomedical Research Center at the			
Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA			
sequencing consortium of the German Genome Project. This clone			
(DKFp313A1611) is available at the RZPD in Berlin. Please contact			
the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059			
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further			

information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cdna/>.

FEATURES

source
1. .3362
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp313M1311"
/tissue_type="human bone marrow"
/clone_lib="313 (synonym: hlcc2). Vector pTriplex2; host
Fobo.G., Han.M. and Wiemann.S.
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
polyA_signal 3308..3313
polyA_site 3332
ORIGIN

Query Match 100.0%; Score 711; DB 9; Length 3362;
Best Local Similarity 100.0%; Pred. No. 4.7e-198;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTCCCTGATCGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCT 60
DB 28 ATGGCGTCCCTGATCGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCT 87
QY 61 TCGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAATCCATCATGACATGTGGC 120
DB 88 TCGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAATCCATCATGACATGTGGC 147
QY 121 TGCTCAGTGGATGTCAGATTCATGATTACAAAGAGAAACCCAGAGAGAGACCTAC 180
DB 148 TGCTCAGTGGATGTCAGATTCATGATTACAAAGAGAAACCCAGAGAGAGACCTAC 207
QY 181 TACATAGAAATATGGGATGTGGAGCTCTGTGGGAGTCCAGCAGCGTGAAAGCACA 240
DB 208 TACATAGAAATATGGGATGTGGAGCTCTGTGGGAGTCCAGCAGCGTGAAAGCACA 267
QY 241 AGAGCAGTATCTACAACTCCGTAAATGGTATTATTTCTGACACGACTTAACAAATAAG 300
DB 268 AGAGCAGTATCTACAACTCCGTAAATGGTATTATTTCTGACACGACTTAACAAATAAG 327
QY 301 AAGTCTCCCAAACTTGGTGTGATTTGAAGTCTCAAGTCTCAACAGGGATTTGGTCCCA 360
DB 328 AAGTCTCCCAAACTTGGTGTGATTTGAAGTCTCAAGTCTCAACAGGGATTTGGTCCCA 387
QY 361 ACTGGAGTCTTGTGCAAAATGGGATTTATGATCAAGAACAGTTTGTGTAACCAATA 420
DB 388 ACTGGAGTCTTGTGCAAAATGGGATTTATGATCAAGAACAGTTTGTGTAACCAATA 447
QY 421 CCACTGTTGTAATAGGACTAACTGGACAGATTATGATCAAGAACAGTTTGTGTAACCAATA 480
DB 448 CCACTGTTGTAATAGGACTAACTGGACAGATTATGATCAAGAACAGTTTGTGTAACCAATA 507
QY 481 TTAAGTGGAGTCTTGTGCAAAATGGGATTTATGATCAAGAACAGTTTGTGTAACCAATA 540
DB 508 TTAAGTGGAGTCTTGTGCAAAATGGGATTTATGATCAAGAACAGTTTGTGTAACCAATA 567
QY 541 ACAAATCCAGGACTTGTGCAAAATGGGATTTATGATCAAGAACAGTTTGTGTAACCAATA 600
DB 568 ACAAATCCAGGACTTGTGCAAAATGGGATTTATGATCAAGAACAGTTTGTGTAACCAATA 627
QY 601 GATAAGTCTATAGAGAGAGATCTTTTAAAGAGAGAGATTAATCAAGTCCAGGCTTTTCT 660
DB 628 GATAAGTCTATAGAGAGAGATCTTTTAAAGAGAGAGATTAATCAAGTCCAGGCTTTTCT 687
QY 661 GATCGGAAGAGATTTGGGCGAGGAGCAATTAAGAGCCTTCATTTATGACTGA 711
DB 688 GATCGGAAGAGATTTGGGCGAGGAGCAATTAAGAGCCTTCATTTATGACTGA 738

RESULT 4

HSR07739 HSR07739 3449 bp mRNA linear PRI 30-AUG-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp313M1333 (from clone DKFZp313M1333).
DEFINITION

ACCESSION

BX647593 GI:34366750

VERSION

BX647593.1

KEYWORDS

Source

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 3449)

AUTHORS

Koehler.K., Beyer.A., Mewes.H.W., Weil.B., Amid.C., Osanger.A.,

Fobo.G., Han.M. and Wiemann.S.

CONSTRM

The German Human cDNA Consortium

JOURNAL

Submitted (27-AUG-2003) MPS, Ingolstaedter Landstr.1, D-85764

COMMENT

Neuberberg, GERMANY

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by BMFZ (Biomedical Research Center at the

Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA

sequencing consortium of the German Genome Project. This clone

(DKFZp313M1333) is available at the RZPD in Berlin. Please contact

the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further

information about the clone and the sequencing project is available

at <http://mips.gsf.de/proj/cdna/>.

FEATURES

source

1. .3449

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp313M1333"

/tissue_type="human heart"

/clone_lib="313 (synonym: hlcc2). Vector pTriplex2; host

DH10B; sites SfiIA + SfiIB"

/dev_stage="adult"

polyA_signal 3311..3316

polyA_site 3336

ORIGIN

Query Match 100.0%; Score 711; DB 9; Length 3449;

Best Local Similarity 100.0%; Pred. No. 4.7e-198;

Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTCCCTGATCGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCT 60

DB 31 ATGGCGTCCCTGATCGGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCT 90

QY 61 TCGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAATCCATCATGACATGTGGGC 120

DB 91 TCGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAATCCATCATGACATGTGGGC 150

QY 121 TGCTCAGTGGATGTCAGATTCATGATTACAAAGAGAAACCCAGAGAGAGACCTAC 180

DB 151 TGCTCAGTGGATGTCAGATTCATGATTACAAAGAGAAACCCAGAGAGAGACCTAC 210

QY 181 TACATAGAAATATGGGATGTGGAGGCTCTGTGGGAGTCCAGCAGCGTGAAAGCACA 240

DB 211 TACATAGAAATATGGGATGTGGAGGCTCTGTGGGAGTCCAGCAGCGTGAAAGCACA 270

QY 241 AGAGCAGTATCTACAACTCCGTAAATGGTATTATTTCTGACACGACTTAACAAATAAG 300

DB 271 AGAGCAGTATCTACAACTCCGTAAATGGTATTATTTCTGACACGACTTAACAAATAAG 330

QY 301 AAGTCTCCCAAACTTGGTGTGATTTGAAGTCTCAAGTCTCAACAGGGATTTGGTGCCA 360

DB 331 AAGTCTCCCAAACTTGGTGTGATTTGAAGTCTCAAGTCTCAACAGGGATTTGGTGCCA 390

QY 361 ACTGGAGTCTTGTGCAAAATGGGATTTATGATCAAGAACAGTTTGTGTAACCAATA 420

DB 391 ACTGGAGTCTTGTGCAAAATGGGATTTATGATCAAGAACAGTTTGTGTAACCAATA 450

QY 421 CCACTGTTGTAATAGGACTAACTGGACAGATTATGATCAAGAACAGTTTGTGTAACCAATA 480

DB 451 CCACTGTTGTAATAGGACTAACTGGACAGATTATGATCAAGAACAGTTTGTGTAACCAATA 510

```

QY 481 TTAACCTAGGACTGCTTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTGGACTGC 540
Db 511 TTAACCTAGGACTGCTTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTGGACTGC 570
QY 541 ACAATCCACGGTACTTAGCTGAGGCTTTCTCAATGCTGTCAAGCTCAGTAGTGTGTTT 600
Db 571 ACAATCCACGGTACTTAGCTGAGGCTTTCTCAATGCTGTCAAGCTCAGTAGTGTGTTT 630
QY 601 GATAAGGTCATAGAGAAGAGATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTTCT 660
Db 631 GATAAGGTCATAGAGAAGAGATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTTCT 690
QY 661 GATCGAAAGATTTGGGCGAGGAAATTAAGAGGCTTCATTATGACTGA 711
Db 691 GATCGAAAGATTTGGGCGAGGAAATTAAGAGGCTTCATTATGACTGA 741

RESULT 5
HSM807810 3352 bp mRNA linear PRI 30-AUG-2003
LOCUS Homo sapiens mRNA; cDNA DKFp3130171 (from clone DKFp3130171).
ACCESSION BX647664
VERSION BX647664.1 GI:34366821
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3352)
AUTHORS Pouetka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
Mewes H.W., Weil B., Amid C., Oeinger A., Fobo G., Han M. and
Wiemann S.
The German Human cDNA Consortium
Direct Submission
Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFp3130171) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

FEATURES
source Location/Qualifiers
1..3352
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFp3130171"
/tissue_type="human fetal skin"
/clone_lib="313 (synonym: hlcc2). Vector pTriplEx2; host
DH10B; sites SfilA + SfilB"
/dev_stage="fetal"
polyA_signal 3296..3301
polyA_site 3321
ORIGIN

Query Match 99.8%; Score 709.4; DB 9; Length 3352;
Best Local Similarity 99.9%; Pred. No. 1.4e-197;
Matches 710; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGTCCCTGGATCGGGTGAAGGTACTGGTGTGGAGACTCAGGTGTGGGAATCT 60
Db 15 ATGGCGTCCCTGGATCGGGTGAAGGTACTGGTGTGGAGACTCAGGTGTGGGAATCT 74
QY 61 TCGTTAGTCCATCTCTATGCCAAATCAAGTCTCGGAAATCCATCATGACTGTGGGC 120
Db 75 TCGTTAGTCCATCTCTATGCCAAATCAAGTCTCGGAAATCCATCATGACTGTGGGC 134

```

```

QY 121 TGCTCAGTGGATGTGAGAGTTCATGATTACAAGAAGAAACCCAGAGAGAGACTTAC 180
Db 135 TGCTCAGTGGATGTGAGAGTTCATGATTACAAGAAGAAACCCAGAGAGAGACTTAC 194
QY 181 TACATAGAAATTAATGGATGTTGGAGGCTCTGTGGGCAAGTCCAGCAGCGTGAAGCACA 240
Db 195 TACATAGAAATTAATGGATGTTGGAGGCTCTGTGGGCAAGTCCAGCAGCGTGAAGCACA 254
QY 241 AGAGCAGTATTTCTACAACTCCGTAATGTAATTTTTCGACAGACTTAACAATAAG 300
Db 255 AGAGCAGTATTTCTACAACTCCGTAATGTAATTTTTCGACAGACTTAACAATAAG 314
QY 301 AAGTCTCTCCAAAACCTTGGTGTGTTGTCATTGGAAGCTCTCAACAGGCAATTTGGTGCCA 360
Db 315 AAGTCTCTCCAAAACCTTGGTGTGTTGTCATTGGAAGCTCTCAACAGGCAATTTGGTGCCA 374
QY 361 ACTGGAGTCTTGGTGACAAATGGGATTAATGATCAAGAAACAGTTTGTCTGATACCAATA 420
Db 375 ACTGGAGTCTTGGTGACAAATGGGATTAATGATCAAGAAACAGTTTGTCTGATACCAATA 434
QY 421 CCAGTGTGTTGTAATAGGACTAAACTGGACCAAGATTTCATGAACAAGGCCCATGAGTT 480
Db 435 CCAGTGTGTTGTAATAGGACTAAACTGGACCAAGATTTCATGAACAAGGCCCATGAGTT 494
QY 481 TTAACTAGGACTGCTTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTGGACTGC 540
Db 495 TTAACTAGGACTGCTTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTGGACTGC 554
QY 541 ACAAAATCCACGGTACTTAGCTGAGGCTTTCTCAATGCTGTCAAGCTCAGTAGTGTGTTT 600
Db 555 ACAAAATCCACGGTACTTAGCTGAGGCTTTCTCAATGCTGTCAAGCTCAGTAGTGTGTTT 614
QY 601 GATAAGTCTATAGAGAAGAGATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTTCT 660
Db 615 GATAAGTCTATAGAGAAGAGATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTTCT 674
QY 661 GATCGAAAGATTTGGGCGAGGAAATTAAGAGGCTTCATTATGACTGA 711
Db 675 GATCGAAAGATTTGGGCGAGGAAATTAAGAGGCTTCATTATGACTGA 725

RESULT 6
BC020832 1999 bp mRNA linear PRI 29-JUN-2004
LOCUS Homo sapiens RAB, member of RAS oncogene family-like 3, mRNA (cdna
clone MGC:23920 IMAGE:471781), complete cds.
ACCESSION BC020832
VERSION BC020832.1 GI:18088571
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1999)
AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins E.S., Wagner E.S., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Sheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,
Carninci P., Frange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullaby S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettman M., Madan A., Rodrigues S.,
Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalski U., Smallos D.E.,
Schnerch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length

```

TITLE

JOURNAL	human and mouse cDNA sequences	Db	190	TACATAGATTATGGGATGTTGGAGGCTCTGTGGGAGTGGCCAGCGTGAAAGCACA	249
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
REFERENCE	2 (bases 1 to 1999)	Qy	241	AGAGCAGTATTCTACAACTCCGTAATATTTTTCGTACACGACTTAAACAAATAAG	300
AUTHORS	Strausberg, R.	Db	250	AGAGCAGTATTCTACAACTCCGTAATATTTTTCGTACACGACTTAAACAAATAAG	309
TITLE	Direct Submission				
JOURNAL	Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	Qy	301	AAGTCCTCCAAAACCTTGGCTGGTTCATTTGGAAGCTCTCAACAGGATTTGGTGCCA	360
		Db	310	AAGTCCTCCAAAACCTTGGCTGGTTCATTTGGAAGCTCTCAACAGGATTTGGTGCCA	369
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov	Qy	361	ACTGGAGTCTTTGGTGACAAATGGGATTTATGATCAAGAACAGTTTGTGTAAACCAATA	420
COMMENT	Contact: MGC help desk				
	Email: csapbs-remail.nih.gov	Db	370	ACTGGAGTCTTTGGTGACAAATGGGATTTATGATCAAGAACAGTTTGTGTAAACCAATA	429
	Tissue Procurement: ATCC				
	cDNA Library Preparation: CLONTECH Laboratories, Inc.	Qy	421	CCACTGTTGTTAATAGGAGCTAAACTCGACAGATTTCATGAAACAAAGCGCCATGAAGTT	480
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)	Db	430	CCACTGTTGTTAATAGGAGCTAAACTCGACAGATTTCATGAAACAAAGCGCCATGAAGTT	489
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305				
	Web site: http://www-shgc.stanford.edu	Qy	481	TTAACTAGGACTGTTTCCCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC	540
	Contact: (Dickson, Mark) mcd@paxil.stanford.edu	Db	490	TTAACTAGGACTGTTTCCCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC	549
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.				
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov	Qy	541	ACAAATCCAGTACTTTAGCTGCAGGTTCTTCCAACTGCTGCAAGCTCAGTAGGTTTTT	600
source	Series: IRAL Plate: 37 Row: m Column: 23	Db	550	ACAAATCCAGTACTTTAGCTGCAGGTTCTTCCAACTGCTGCAAGCTCAGTAGGTTTTT	609
	This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 28376659.	Qy	601	GATAAGTCTATAGAGAAGAGATATCTTTTAAAGAGAAGTAAATCAGATTTCCAGGCTTCTCT	660
	Location/Qualifiers	Db	610	GATAAGTCTATAGAGAAGAGATATCTTTTAAAGAGAAGTAAATCAGATTTCCAGGCTTCTCT	669
	1. .1999				
	/organism="Homo sapiens"	Qy	661	GATCGGAAAAGATTTGGGCGAGGAACATTAAGAGCCTTCATTTAGACTGA	711
	/mol_type="mRNA"	Db	670	GATCGGAAAAGATTTGGGCGAGGAACATTAAGAGCCTTCATTTAGACTGA	720
	/db_xref="taxon:9606"				
	/clone="MGC:23920 IMAGE:4771781"				
	/tissue_type="Testis, embryonal carcinoma"				
	/clone_lib="NIH_MGC_61"				
	/lab_host="DH10B"				
	/notes="vector: pDNR-LIB"				
gene	1. .1999				
	/gene="RABL3"				
	/notes="synonym: MGC23920"				
	/db_xref="LocusID:285282"				
CDS	10. .720				
	/gene="RABL3"				
	/codon_start=1				
	/product="RAB, member of RAS oncogene family-like 3"				
	/protein_id="AAH20832.1"				
	/db_xref="GI:18088572"				
	/db_xref="LocusID:285282"				
	/translation="MASLDRVKVLVLGDSGVGKSLVHLLCQNVLGNPSWTVCSSVD VRVHYKEGTPPEKTCYIELWDVGSVGSASVSKSTRVAFVNSVNGIIFVHDLTNKKS SQRLRWSLRALNRDLVPTGVLVTNGYDQEPADNOIPLVIGTKLDQIHTKREH LTTATFLAEFDPFEEINLCTNPRYLAAGSSNAVKLSRFPDKVIERKYPFLREGNQIPG PPRKRFAGATLSLHYD"				
ORIGIN					
Query Match	99.5%; Score 707.8; DB 9; Length 1999;				
Best Local Similarity	99.7%; Pred. No. 4e-197;				
Matches	709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
Qy	1 ATGGCGTCCCTGGATCGGTTGAAGGTACTGGTGTGGAGACTCAGGTGTTGGAAATCT	60			
Db	10 ATGGCGTCCCTGGATCGGTTGAAGGTACTGGTGTGGAGACTCAGGTGTTGGAAATCT	69			
Qy	61 TCGTTAGTCCATCTCCTATGCAAAATCAAGTGTGGGAATCCATCATGACTGTGGC	120			
Db	70 TCGTTAGTCCATCTCCTATGCAAAATCAAGTGTGGGAATCCATCATGACTGTGGC	129			
Qy	121 TGCTCAGTGGATGTCAGAGTTCATGATTACAAAGAGAACCCAGAGAACACCTAC	180			
Db	130 TGCTCAGTGGATGTCAGAGTTCATGATTACAAAGAGAACCCAGAGAACACCTAC	189			
Qy	181 TACATAGATTATGGGATTTGGAGGCTCTGTGGGAGTGGCCAGCGTGAAAGCACA	240			

RESULT 7

BC050194

LOCUS

DEFINITION

Mus musculus RAB, member of RAS oncogene family-like 3, mRNA (cDNA

clone MGC:60441 IMAGE:30022531), complete cds.

BC050194

VERSION

BC050194.1 GI:29747934

KEYWORDS

MGC.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2064)

REFERENCE

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Villalón, D.C., Hales, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shavchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,

Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

MEDLINE

22388257

12477932

REFERENCE

2 (bases 1 to 2064)

Strausberg, R.
 Direct Submission
 Submitted (02-APR-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Janet Rossant and Tilo Kunath (Samuel
 Lunenfeld Research Institute, Canada)
 cDNA Library Preparation: Yulan Piao and Minoru Ko (National
 Institute on Aging, NIH: <http://lgsun.grc.nia.nih.gov/cDNA/>)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
 Kowib, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 111 Row: m Column: 3
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13385791.
 Location/Qualifiers
 1. .2064
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="B5/EGFP transgenic ICR mice"
 /db_xref="taxon:10090"
 /clone="MGC:60441 IMAGE:30022531"
 /tissue type="Trophoblast Stem Cell, mouse, 3.5 dpc"
 /clone_lib="NIA Mouse Trophoblast Stem Cell cDNA Library
 (Long)"
 /lab_host="DH10B"
 /notes="Vector: pSPORT1"
 1. .2064
 /gene="Rab13"
 /db_xref="LocusID:67657"
 /db_xref="MGI:1914907"
 3. .713
 /gene="Rab13"
 /codon_start=1
 /product="Rab13 protein"
 /protein_id="AAH50194.1"
 /db_xref="GI:29747935"
 /db_xref="LocusID:67657"
 /db_xref="MGI:1914907"
 /translations="MASLDRVKVLVLGDSGVGKSSLVHLLCHNQLVGNPSWTVGCSVD
 IRVHDKGTEPEKTYVIELVDVGSGVSGASSVKSTRVFNYSVNGIILVHDLTNKKS
 SONLYEWSLEVLNRAVPTGLVTNGDYDQFADNQIPLLVIGTKLDIOLHETKREH
 LRTAFLEADFNABEINLDCNTNPRSSAAGSSNAVKLSRPFDPKVIKRYFPREGNQIPG
 LRSKRPFGGALKNFHCD"
 21. .464
 /gene="Rab13"
 /notes="RAB; Region: Rab subfamily of small GTPases"
 /db_xref="CDD:smart00175"
 misc_feature
 82.7%; Score 587.8; DB 10; Length 2064;
 Best Local Similarity 89.2%; Pred. No. 9e-162;
 Matches 634; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
 1 ATGGCGTCCCTGGATCGGTTGAGGTACTGTTGGAGACTCAGGTGTTGGAAATCT 60
 |||
 3 ATGGCGTCCCTGGATCGGTTGAGGTACTGTTGGAGACTCAGGTGTTGGAAATCT 62
 |||
 61 TCCTTAGTCCATCTCTATCCAAATCAAGTGTGGAAATCCATCATGACTGTGGGC 120
 |||

AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT
 63 TCACTCGTCCATCTTCTGTGCCACAATCAAGTGTGTAGGAAATCCGTCATGGACTGTGGGC 122
 121 TGCTCAGTGGATGTCAGAGTTCATGATTACAAGAAGGACCCAGAGAGAGACTTAC 180
 |||
 123 TGCTCGGTGGATATCAGAGTTCATGACTACAAGAAGGACCCCTGAAGAGAGACATAC 182
 |||
 181 TACATAGAAATATGGGATGTTGGAGGCTCTGTGGGCGAGTCCAGCAGCGTGAAGAGCACA 240
 |||
 183 TATATAGAACTATGGATGTTGGAGGCTCTGTGGGCGAGTCCAGCAGTGTGAAGAGCACA 242
 241 AGAGCAGTATTTTACAACTCCGTAATGTTATTTTCGTACACGACTTAACAAATAAG 300
 243 AGAGCGGTCTTCTACAACTCTGTAAACGGCATCATTTTAGTAGTACATGACTTAACAAATAAG 302
 301 AGTCTCTCCAAACTTGGGTGTTGGTCAATTTGGAGCTCTCAACAGGAGTATTTGGTGCCA 360
 |||
 303 AAGTCATCTCAAAACTTATATCGTGGTCAATTTGGAGTTCCTCAATCGGATGCGATTCCA 362
 361 ACTGGAGTCTTGGTGACAAATGGGATTTATGATCAAGAACAGTTTCTGATTAACCAAATA 420
 363 ACCGAGTCTCTGTGACAAATGGGATTTATGACCGGAACAGTTTCTGATTAACCAAATC 422
 421 CCACTGTTGGTAATAGGGACTAAACTGGACCAAGATTCATGAACAAAGCGCCATGAAGTT 480
 423 CCACTGTTGGTAATAGGGACTAAACTGGACCAAGATTCATGAACAAAGCGCCATGAAGTT 482
 481 TTAACCTAGAGCTCTTCTGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 540
 483 TTAATCAGGACTGCTTCTGCTGAGGACTTCAATGCTGGAAGAGATTAAATTTGGATTGC 542
 541 ACAATCCACGCTACTTAGCTGCAGGTTCTTCAATGCTGTCAAGCTCAGTAGTATTTTT 600
 543 ACAATCCACGCTCTCAGCTGCAGGTTCTTCAATGCTGTCAAGCTCAGTAGTATTTTT 602
 601 GATAAGTTCATAGAGAAGATACCTTTTAAAGAGAAGGTAATCAGATTCCAGCTTTTCCT 660
 603 GATAAGTTATAGAGAAGATATTTTTTTAGAGAAGGTAATCAGATTCCGGGCTTTTCT 662
 661 GATCGGAAAGATTTGGGCGAGGACATTAAGAGGCTTCATATGACTGA 711
 |||
 663 GATCGGAGAGGTTTGGAGGGGCACTGAAGAACTTCCACTGTGACTGA 713
 RESULT 8
 AJ720020 3327 bp mRNA linear VRT 30-SEP-2004
 LOCUS
 DEFINITION
 Gallus gallus mRNA for hypothetical protein, clone 9j2.
 ACCESSION
 AJ720020
 VERSION
 AJ720020.1 GI:53130699
 KEYWORDS
 ORF1.
 SOURCE
 Gallus gallus (chicken)
 ORGANISM
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE
 1
 AUTHORS
 Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezzubov, Y., Zaim, J.,
 Fiedler, P., Kutter, S., Blagodatski, A., Kostovska, D., Koter, M.,
 Plachy, J., Carninci, P., Hayashizaki, Y. and Buerstedde, J.M.
 TITLE
 Full-length cDNAs from bursal lymphocytes to facilitate gene
 function analysis
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 3327)
 AUTHORS
 Caldwell, R.B.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
 Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
 1, D-85764 Neuberg, GERMANY
 FEATURES
 source
 1. .3327
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="CB"

```
/db_xref="taxon:9031"
/clone="9j2"
/cell_type="lymphocyte"
/tissue_type="bursa"
/clone_lib="riken1"
/dev_stage="2 weeks old"
1..3327
/locus_tag="RCJMB04_9j2"
<1..30
/locus_tag="RCJMB04_9j2"
31..723
/locus_tag="RCJMB04_9j2"
/note="ORF1"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAG31679.1"
/db_xref="GI:53130700"
/tranlations="MAALDRVKLVLDGSGVSKSLVHLCHNOVLGNPSWTVGCSVD
VRIDYKEGTPPEKTYIYELMDVGGVSGSVTSKSTRVFNYSINGLIDILVHDLTNKKS
SQLYXWSLEALNDVAPTGLVNTGDIYDREQFADNQIPLLVIGTKLDQILPETHRKEV
LRTAFLEDFNAEEINLCTNPRYLAAGSSNAVLKLSRPFDKVIEKRYFPRWQSRFL
ASRKERDLBEEH"
724..>3327
/locus_tag="RCJMB04_9j2"

ORIGIN
Query Match      68.1%; Score 484.2; DB 5; Length 3327;
Best Local Similarity 81.8%; Pred. No. 3.1e-131;
Matches 583; Conservative 0; Mismatches 128; Indels 2; Gaps 2;

QY 1 ATGGCGTCCGTGATCGGTTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAATCT 60
DB 1 ATGGCGGCTCTGGATAGGGTCAAGGTGCTGGTGTGGGCGACTCCGGTGTGGGAATCT 60
QY 61 TCGTTAGTCCATCTCTATCCAAAATCAAGTGTGGGAATCCATCATGACGTGTGGGC 120
DB 61 TCGTCTGTGACCTCTCTGTGTCAACAGGTGTGGGAACCCGTCTGGAGCGTGGGC 150
QY 121 TGCTCAGTGTGATCTCAGATTTCATGATTACAAAGAGAAACCCAGAGAGAACCTAC 180
DB 151 TGCTCGTGTGATGCGAATCCATGACTACAAAGAGGGACTCGGAAGAGAGACATAC 210
QY 181 TACATAGATTTGGATGTGGAGGCTGTGGGCGAGTCCGACGAGCGTGAAGAAGACA 240
DB 211 TATATAGAGCTGTGGATGTGGAGGTTCAGTGGGTAGTGCCACTAGTGTGAAGAAGACA 270
QY 241 AGAGCAGTATCTACAACTCGTAAATGATTTATTTTCCTACAGCTTAAACAATAAG 300
DB 271 AGAGCAGTATTTTATAACTCGTGAATGGGATAATTTTGTGATGACTTAAACCAAG 330
QY 301 AAGTCCCTCCAAAATTTGCGTGTGGTCAATTTGAGAGCTCTCAACAGGGATTTGGTGCCA 360
DB 331 AATATATCCAGATTTATATCGTGTCTTTTGAAGACTCAACAGAGATGTGCTCCC 390
QY 361 ACTGGAGTCTTGTGACAAATGGGATTTATGATCAAGAACAGTTTGTCTGATTAACCAATA 420
DB 391 ACAGGAGTCTTGTGACAAACGGTGTACTACGCGTGAAACAGTTTGTCTGATTAACAGATT 450
QY 421 CCAGTGTGTGATAGGACTAACTGGACAGATTTCATGAACAAAGCCGATGAAGTT 480
DB 451 CCAGTGTGTGATAGGACTAACTGGATGATTCAGATCCAGAAACAAAGCGTAAATGAAGTT 510
QY 481 TTAAGTGGAGTCTTTCCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGAGTGC 540
DB 511 TTGACAGGACAGCTTTCTTGGCTGAGGATTTCAATGCTGAGAGATTAATTTGGATTGC 570
QY 541 ACAAAATCCAGGTTACTAGTGCAGGTTCTTCCAAATGCTGTCAAGCTCAGTAGTTTTTT 600
DB 571 ACCAATCTCTGTACTTGGCTGAGGCTCATCCAAATGCTGTCAAACTAAGCAGATTTTT 630
QY 601 GATAAGTGTATAGAGAGAGATACCTT-TTTAAGAGAGAGGTAATC-AGATTCCAGGCTTTC 658
DB 631 GATAAGTGTATAGAGAGAGATACCTTCTTTAAGAGAGATGGCAATCAAGATTCTCTGGCTTCT 690

QY 659 CTGATCGGAAGAGATTGGGCGAGGAACATTAAAGAGCCCTTCATTATGACTGA 711
DB 691 CGGAAGAAAGAGATTGGGAGGAGGAACACTGAAGAGCCCTTCATTATGACTGA 743

RESULT 9
AK025772 2550 bp mRNA linear PRI 13-SEP-2003
LOCUS Homo sapiens cDNA: FLJ222119 fis, clone HEP18852.
DEFINITION AK025772
ACCESSION AK025772
VERSION 1 GI:10438390
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE 1
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2550)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail: flicdn@ims.u-tokyo.ac.jp, tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
location/Qualifiers
1..2550
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP18852"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/notes="cloning vector pME18SFL3"

ORIGIN
Query Match      61.2%; Score 435.4; DB 9; Length 2550;
Best Local Similarity 99.8%; Pred. No. 7.2e-117;
Matches 436; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 275 TTTTCGTACACGACTTAAACAAATAAGAGTCCTCCCAAACTTGCCTGTCATTGG 334
DB 1 TTTTCGTACACGACTTAAACAAATAAGAGTCCTCCCAAACTTGCCTGTCATTGG 60
QY 335 AAGCTCTCAACAGGATTTGGTCCCAACTGGAGTCTTGGTGACAAATGGGATTTATGATC 394
DB 61 AAGCTCTCAACAGGATTTGGTCCCAACTGGAGTCTTGGTGACAAATGGGATTTATGATC 120
QY 395 AAGCAACAGTTTGTGTGATACCAATACCACTGTTGGTAATAGGACTAAACTGACACAGA 454
DB 121 AAGCAACAGTTTGTGTGATACCAATACCACTGTTGGTAATAGGACTAAACTGACACAGA 180
QY 455 TTCAATGAAAACAAAGCGCCATGAAGTTTAACTAGGACTGCTTCTCGCTGAGGATTCA 514
DB 181 TTCAATGAAAACAAAGCGCCATGAAGTTTAACTAGGACTGCTTCTCGCTGAGGATTCA 240
QY 515 ATCCAGAGAAATTAATTTGGAGCTGCAAAATCCACGGTACTTAGCTGCGAGGTTCTTCCA 574
DB 181 ATCCAGAGAAATTAATTTGGAGCTGCAAAATCCACGGTACTTAGCTGCGAGGTTCTTCCA 574
```

Db	241	ATCCAGAGAAATTAATTTGGACTGCACAAATCCCGGTACTTAGCTGCAGGTTCTTCCA	300
QY	575	ATGCTGTCAAGCTCAGTAGGTTTTTGTATAGTCTATAGAGAGATATCTTTTAAAGAG	634
Db	301	ATGCTGTCAAGCTCAGTAGGTTTTTGTATAGTCTATAGAGAGATATCTTTTAAAGAG	360
QY	635	AAGGTAAATCAGATTCACAGGCTTCCCTGATCGGAAAGATTTGGGGCAGGAACATTAAGA	694
Db	361	AGGTAAATCAGATTCACAGGCTTCCCTGATCGGAAAGATTTGGGGCAGGAACATTAAGA	420
QY	695	GCCTTCATTATGACTGA 711	
Db	421	GCCTTCATTATGACTGA 437	
RESULT 10			
BC073035			
LOCUS			
DEFINITION	BC073035	1064 bp mRNA linear VRT 03-AUG-2004	
		Xenopus laevis MGC82648 protein, mRNA (cDNA clone MGC:82648	
		IMAGE:5049322), complete cds.	
ACCESSION	BC073035		
VERSION	BC073035.1	GI:49257300	
KEYWORDS			
SOURCE			
ORGANISM			
		Xenopus laevis (African clawed frog)	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;	
		Xenopodinae; Xenopus; Xenopus.	
REFERENCE			
AUTHORS		1 (bases 1 to 1064)	
		Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.	
		and Richardson, P.	
TITLE		Genetic and genomic tools for Xenopus research: The NIH Xenopus	
		initiative	
JOURNAL		Dev. Dyn. 225 (4), 384-391 (2002)	
PUBMED		12454917	
REFERENCE			
AUTHORS		2 (bases 1 to 1064)	
		Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,	
		Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,	
		Altschul, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,	
		Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,	
		Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,	
		Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,	
		Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,	
		Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,	
		Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,	
		McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,	
		Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,	
		Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,	
		Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,	
		Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,	
		Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,	
		Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,	
		Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,	
		Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.	
		Generation and initial analysis of more than 15,000 full-length	
		human and mouse cDNA sequences	
TITLE			
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
PUBMED		12477932	
REFERENCE			
AUTHORS		3 (bases 1 to 1064)	
		Klein, S. and Gerhard, D.S.	
TITLE		Direct Submission	
JOURNAL		Submitted (07-JUN-2004) National Institutes of Health, Xenopus Gene	
		Collection (XGC), National Institute of Child Health and Human	
		Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD	
		20892-7510, USA	
REMARK			
COMMENT		NIH-MGC Project	
		Contact: XGC help desk	
		Email: cgaaps-r@mail.nih.gov	
		Tissue Procurement: Dr. Igor David	
		cDNA Library Preparation: Life Technologies, Inc.	
		DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)	
		BC Cancer Agency, Vancouver, BC, Canada	

info@bcgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu, Parvaneh Saeeidi, JR Santos, Angeliue Schnerch, Ursula Skalska, Duane Smallos, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 152 Row: d Column: 10

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES	Location/Qualifiers
source	1..1064
	/organism="Xenopus laevis"
	/mol_type="mRNA"
	/db_xref="taxon:8355"
	/clone="MGC:82648 IMAGE:5049322"
	/tissue_type="Ovary, Xenopus"
	/clone_lib="NICHD_XGC_Ov1"
	/lab_host="DH10B"
	/note="Vector: pCMV-SPORT6"
gene	1..1064
	/gene="MGC82648"
	/db_xref="LocusID:444046"
	35..742
	/gene="MGC82648"
	/codon_start=1
	/product="MGC82648 protein"
	/protein_id="AAH73035.1"
	/db_xref="GI:49257301"
	/db_xref="LocusID:444046"
	/translation="MASLDKRVKVLVLDGSGVSKSLVHLLCONVLGNPSWTGVCSDV
	VLHVRSGTPEKTYTTELWDGSGVSGSASSVKSTRAVYNAVNGILVHDLTNKKS
	SNLRYRSLBALNRLOPMGVLTGDDYDQFADNQIPLLVIGTGLDQIPEAKRNEV
	LRTFLAFLEDFNAEEINLDCNTNRCLAGSSNAVKLSRFDFKRIEKPREGNLIIPGF
	SDRKRFPGGNGFKSLHYD"

ORIGIN	
Query Match	58.8%; Score 418.2; DB 5; Length 1064;
Best Local Similarity	75.2%; Pred. No. 8.1e-112;
Matches 535; Conservative	0; Mismatches 173; Indels 3; Gaps 1;
QY	1 ATGGCGCTCCCTCGATCGCGTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAATCT 60
Db	35 ATGGCTTCTCTGGACAGAGTGAAGGTTCTGGTCTGGGGGATTCAGGAGTCGGGAAGTCT 94
QY	61 TGTGTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGGACTGTGGGC 120
Db	95 TCGTTGGTGCATCTGCTGTGTGAGATATCAGGTTTTCGGGAAACCCCTCGTGGACAGTTGGC 154
QY	121 TGTCTAGTGTGATGTGAGTTCATGATTACAAGAAGAGAACCCAGAGAGAGACCTAC 180
Db	155 TGCTCAGTGATGTGCGGCTCCACGATACAGGAAGGAGACTCCAGAGGAAACCTAT 214
QY	181 TACATAGAAATATGGGATGTGGAGGCTCTGTGGCAGTCCAGCAGCGGTGAAGACACA 240
Db	215 TACACTGAGCTGTGGGACGTGGGTCCGTTGGCAGTCCAGCAGTGTGAAGAGCACC 274
QY	241 AGAGCAGTATTTCTACAACTCCGTAATGTTATTTTGTGACACAGCTTAACAATAAG 300
Db	275 AGGGCTGTGTTTATTAATCAGTAAATGGATCATCTTTGTCATGATTCACAAAATAAG 334
QY	301 AAGTCTCTCCAAAACCTTGGTCTGTTGCTCATTTGGAAGCTCTCAACAGGAGTTTGGGCCA 360
Db	335 AAATCATCTCAGAACCTTTTATCGTGTCTATTAGAACCCCTGAACCGAGATCTGCAGCCA 394
QY	361 ACTGGAGTCTTGGTGAACAAATGGGGATTATGATCAAGAACAGTTTGTCTGATTAACCAATA 420

Db 395 ATGGAGTGTAGTCACTAATGGGACTACAGACAGAGAGCAATTTGCTGATATCAGATC 454

Qy 421 CCACTGTGTGTAATAGGAGCTAAATCTGACAGATTCATGAACAAAGCGCCATGAAGTT 480

Db 455 CCATTGTGTGTCATTTGGAACCAAACTGACCAAAATCCAGAGCGCAAGCAATGAGGTC 514

Qy 481 TTAACCTAGGACTGTTTCCCTGGCTGAGGATTCATCCAGAGCAATTAATTTGGACTGC 540

Db 515 CTGACACGAAGTCCCTTTCTAGCTGAGGATTTTAATGCTGAAGAGATTAATCTTGACTGT 574

Qy 541 ACAAAATCCAGCTACTTAGCTGAGGTTCTTCCAAATCTGCTCAAGCTCAGTAGTGTTCCT 600

Db 575 ACAGACACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634

Qy 601 GATAAGTCTATAGAGAGATGATCTTTTAAAGAGAGGTAATCAGATTCAGGCTTTTCCT 660

Db 635 GACAGAGTTTATGAGAAAGATAC---CCACGGGAGGAAATCTGATCCCTGGTTTCTCC 691

Qy 661 GATCGGAAGATTTGGGCGAGGACATTAAGAGCCTTCATTTATGACTGA 711

Db 692 GATCGGAACGCTTTGGAGGGGGAGCTTTAAGAGTCTCCACTATGACTGA 742

RESULT 11

BC078191

LOCUS BC078191 1040 bp mRNA linear VRT 26-JUL-2004

DEFINITION Danio rerio cDNA clone MGC:100876 IMAGE:7141370, complete cds.

ACCESSION BC078191

VERSION BC078191.1 GI:50603855

KEYWORDS MGC.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

REFERENCE 1 (bases 1 to 1040)

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Shennan, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hong, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullighy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.B., Jones, S.J., and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 1040)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (19-JUL-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk
Email: cgabs-x@mail.nih.gov
Tissue Procurement: Genome Institute of Singapore
cDNA Library Preparation: S. Mathavan, Chia-Lin Wei, and Yijun Ruan, Genome Institute of Singapore
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Heshighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
Series: IRL Plate: 60 Row: e Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source 1.1040

/organism="Danio rerio"

/mol_type="mRNA"

/strain="Singapore local strain"

/db_xref="taxon:7955"

/clone="MGC:100876 IMAGE:7141370"

/tissue_type="Embryo, 7 different stages (from just fertilized embryos to 72 hours just hatched baby fish)"

/clone_lib="GIS2F001_ra"

/lab_host="DH10B"

/note="Vector: pDNR-LIB"

42..743

/codon_start=1

/product="Unknown (protein for MGC:100876)"

/protein_id="AAH78191.1"

/db_xref="GI:50603855"

/translation="MASLDRVKVLVLDGSGVKSLLVHLLCQVNLGNPSWTVCSDVVRVHDREGTPEEKAFYIELVDVGVSGSVKSTRVFNYSVNGILLVHDLTKKVSQNLVLEALSKDSPTGITVNSGDYDREQFAENAVPLLLITGTFDQIPENKRNVDLRTAFSLSDFNABEINLDCNTNPRYLAAGSSNAVKLSRFDFKVIKRYFTDPDSQMSFTDRRRFNFKSLHSD"

CDS

Query Match 55.8%; Score 396.6; DB 5; Length 1040;
Best Local Similarity 74.2%; Pred. No. 1.9e-105;
Matches 501; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

ORIGIN

Qy 1 ATGGCGTCCCTCGATCGGTGAAGTACTGGTCTTGGAGACTCAGGTGTGGGAATCT 60
Db 42 ATGGCATCTTTGGACAGGGTGAAGTCTGGTCTTGGAGATCTTGGAGTAGGGAATCC 101
Qy 61 TCCTTAGTCCATCTCTATGCCAAAATCAAGTCTGGAAATCCATCATGCACTGTGGGC 120
Db 102 TCCTTTGTACACCTCTTTTGGCAGATCAGTTTGGGGATCCCTCATGCACTGTGGGC 161
Qy 121 TGCTCAGTGGATGTCAAGTTCATGATTACAAAGAGAAACCCAGAGAGAACCTTAC 180
Db 162 TGCTCAGTGGACGTCAAGGTTTCAAGTACAGAGAGGCACTCCAGAGAGAGGCTTTC 221
Qy 181 TACATAGATTTATGGGATGTGGAGCTCTGTGGGAGTCCAGCAGCTGGAAGACACA 240
Db 222 TACATTTGAATCTGGGATGTGGAGGATCTGTGGCAGTCCAGCAGTGTGTAAGGACCC 281
Qy 241 AGAGCAGTATTTACAACTCCGTAAATGGTATTATTTCTGACAGCACTTAAACAAATAG 300
Db 282 AGAGCTGTGTTTACAAATTCAGTTAATGGTATCATTTTAGTTCAGTCTGACTACAG 341
Qy 301 AAGTCTCTCCAAACTTGGCTCGTTCATTTGGAGTCTCAACAGGGATTTGGTCCCA 360
Db 342 AAGTCTCTCCAGAACTTATACCGCTGTCTATTAGAGGCACTAAGCAGGAGCTCTCTCCA 401
Qy 361 ACTGAGTCTTGTGTGCAAAATGGGATTTATGATCAAGACAGTTTGTGTATACCAATA 420

Db	402	ACTGGCATCATCGPATCAACGGGTGATTCAGATAGAGAACAGTTTTCAGAGAAACGCTGTT	461
Qy	421	CCACTGTGTGTAATAGGAGCTAAACTCGACACAGATTTCATGAAACAAAGCGCATGAAGTT	480
Db	462	CCTCTTCTGCTAATTTGGCACCCTAATTTGATCAGATCCGGGAAACAGAGGACGATGTC	521
Qy	481	TTAACTAGGACTGCTTTCTCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGAGCTGC	540
Db	522	CTGACCCGAACTGCTCTTCTATCTGGAAGACTTCAATGCGGAAGAGATCAACCTGGATTGT	581
Qy	541	ACAAATCCAGGTACTTAGCTGCGAGTCTTCCAAATGCTGTCAGCTCAGTAGGCTTTTTT	600
Db	582	ACAAATCCAGGTATCTTGCTGCTGGTTTCATCAAAATGCTGTTAAATTTGAGTCGATCTTT	641
Qy	601	GATAAGGCTCATAGAGAGAGATACCTTTTTTAAGAGAAGTAAATCAGATTCAGGCTTTTCC	660
Db	642	GATGAGGTATAGAGAAAGATATTTTACAAGAGACCTTAGCCAGATGCAGAGCTTCACA	701
Qy	661	GATCGGAAAGATTT	675
Db	702	GACAGGAGGCGCTTT	716
RESULT 12			
LOCUS	AY391440	1291 bp	mRNA linear VRT 07-OCT-2003
DEFINITION	Danio rerio hypothetical protein MGC23920-like protein mRNA, complete cds.		
ACCESSION	AY391440		
VERSION	AY391440.1	GI:37362247	
KEYWORDS			
SOURCE	Danio rerio (zebrafish)		
ORGANISM	Danio rerio		
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.		
AUTHORS	1 (bases 1 to 1291)		
TITLE	Song,H.D., Wu,X.Y., Sun,X.J., Zhou,Y., Liu,T.X., Deng,M., Zhang,G.W., Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y., Zou,L.I., Kanki,J.P., Look,A.T. and Chen,Z.		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 1291)		
TITLE	Song,H.D., Wu,X.Y., Sun,X.J., Zhou,Y., Liu,T.X., Deng,M., Zhang,G.W., Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y., Zou,L.I., Kanki,J.P., Look,A.T. and Chen,Z.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (15-SEP-2003) State Key Lab for Medical Genomics, Shanghai Institute of Hematology, Ruijin Hospital Affiliated to Shanghai Second Medical University, 197 Rui Jin Road II, Shanghai 200025, P. R. China		
TITLE	Location/Qualifiers		
FEATURES	1. 1291		
source	/organism="Danio rerio"		
	/mol_type="mRNA"		
	/db_xref="taxon:7955"		
	/clone="RK025A2D04"		
	3. 704		
	/codon_start=1		
	/product="hypothetical protein MGC23920-like protein"		
	/protein_id="AA031252.1"		
	/db_xref="GI:37362248"		
	/translation="MASDIRVKVLVLGDSGVGKSSLVLLHCQNVLPNSWTVGCSVDVRVHDYRGEPEEKAFYELWDVGSGVSSVSKSTRVFNYSVLLVHDLTNKKSQNLVRSGLSKSDSPGTGIIVSNGDYDREQFNAVPLLLGTGKPDQIPENKRNVDLRTAFILSDFNABEINLDCNTNPRYLAAGSSNAVKLSRFDFKVIETKRYFTRDPSQMSFTDRLRRNFNFKSLHSD"		
CDS			
ORIGIN			
Query Match	55.6%	Score 395;	DB 5; Length 1291;
Best Local Similarity	74.1%;	Pred. No.5.6e-105;	
Matches 500; Conservative	0;	Mismatches 175;	Indels 0; Gaps 0;

QY	1	ATGCGCTCCTGATCGGGTGAAGGTA	CTGGTGTGGGAGACTCAGGTGTGGGAAATCT	60
Db	3	ATGGCATCTTTGGCAGGGTGAAGTGT	CTGGTGTGGGAGATTCGAGTAGGGAAGTCC	62
QY	61	TCGTTAGTCCATCTCTATGCGGAAAT	CAAGTCTCTGGGAAATCATCATGAGCTGTGGGC	120
Db	63	TCTCTTGTACACCTTCTTTGCGGAA	TCAGTGTGGGAAATCCCTCATGGACTGTGGGC	122
QY	121	TGCTCAGTGTGATGTGAGAGTTCAAT	GATTCAAAAGAGAGGAACCCAGAGAGAGAGAGACTAC	180
Db	123	TGCTCAGTGTGAGTGTGAGGTTTCA	GACTACAGAGAGGCACTCTCAGAGAGAGAGGCTTTC	182
QY	181	TACATAGAAATATGAGGATGTGGAG	GCTCTGTGGCAGTCCAGCAGCGGTGAAAAGCACA	240
Db	183	TACATTTGAACTCTGGGATGTGGAG	GATCTGTGGCAGTCCAGCAGCTGTAAAAGCACC	242
QY	241	AGAGCAGTATTTACAACTCCGTAAT	TGGTATTTATTTTCGTACACGACTTAAACAATAAG	300
Db	243	AGAGCTGTGTTTACAAATTCAGTTA	TGTTAGTTCACGATCTCGACTAACAAG	302
QY	301	AAGTCTCCCAAACTTGGCTGTGGT	CAATTCGGAAGCTCTCAACAGGGAATTTGGTGCCA	360
Db	303	AAGTCTCCCAAGATCTATACCGCT	TGCTACTAGAGCACTGAGCAAGGACTCCTCTTCCA	362
QY	361	ACTGGAGTCTTGGTGAACAATGGGG	ATATGATCAAGAAACAGTTTGTCTGATACCAAAAT	420
Db	363	ACTGGCATCATGTATCCAACTGAT	TACGATAGAGAAACAGTTTGCAGAGAAATGCTGTT	422
QY	421	CCACTGTTGGTATATAGGGACTAA	ACTGGACAGATTCATGAACAAAGCCCAATGAAGTT	480
Db	423	CCTCTTCTGTAAATGGCACAAAT	TTGATCAGATCCCGGAAACAGAGGAACGATGTC	482
QY	481	TTAACTAGGATGCTTTCCTGGCTG	AGGATTTCAATCCAGAGAAATTAATTTGGAGTGC	540
Db	483	CTGACCCGAAGTGCCTTCTTATC	TGAAGACTTCAATCGGGAAGAGATCAACCTGGATTC	542
QY	541	ACAAATCCAGGTAATAGCTGCAAG	TTCTTCAATGCTCTCAAGTTCAGTAGGTTTTTTT	600
Db	543	ACAAATCCAGGTAATCTTGTCTG	TGTTTCAATCAATGCTGTAAACTGAGTTCGATCTTT	602
QY	601	GATAAGTCTATAGAGAGAGATACT	TTTTTAAGAGAGAGGTAATCAGATTCAGGCTTTCT	660
Db	603	GATAAGGTAATAGAGAAAGATACT	TTCACAAAGAGACCCCTAGCCAGATGCAGAGCTTCACA	662
QY	661	GATCGGAAAGATTT	675	
Db	663	GACAGGAGCGCTTT	677	
RESULT	13			
BD076273		BD076273	385 bp	DNA linear
LOCUS				PAT 27-AUG-2002
DEFINITION				5' EST of tissue-nonspecific secretory protein.
ACCESSION				BD076273
VERSION				BD076273.1 GI:22621876
KEYWORDS				JP 2001512011-A/221.
SOURCE				Homo sapiens (human)
ORGANISM				Homo sapiens
REFERENCE				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS				1 (bases 1 to 385)
TITLE				Edwards, J.B.D.M., Duclert, A. and Lacroix, B.
JOURNAL				5' EST of tissue-nonspecific secretory protein
COMMENT				Patent: JP 2001512011-A 221 21-AUG-2001;
				GENSET
				OS Homo sapiens (human)
				PN JP 2001512011-A/221
				PD 21-AUG-2001
				PF 31-JUL-1998 JP 2000505289
				PR 01-AUG-1997 US 08/905135
				PI JEAN BAPTISTE DUMAS MILNE EDWARDS, AYMERIC DUCLERT, BRUNO PI LACROIX
				PC C12N15/09, C12N15/09, C07K14/47, C12Q1/68, C12N15/00, C12N15/00 CC

```
blastn
CC identity 97
CC region 1..347
CC id AA023764
CC est
CC blaetn
CC identity 95
CC region 145..384
CC id C03036
CC est
CC blaetn
CC identity 98
CC region 2..71
CC id C03036
CC est
CC blaetn
CC identity 99
CC region 1..193
CC id R08519
CC est
CC blaetn
CC identity 94
CC region 193..263
CC id R08519
CC est
CC Von Heijne matrix
CC score 4.8
CC seq SLVHLCONQVLG/NP
CC n=a, g, c or t
FH Key Location/Qualifiers
FT misc_feature 39..385
FT misc_feature 146..385
FT misc_feature 11..80
FT misc_feature 39..231
FT misc_feature 232..302
FT sig_peptide 11..109
FT misc_feature 287
FT misc_feature 331.
FEATURES
source
1..385
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 51.4%; Score 365.8; DB 6; Length 385;
Best Local Similarity 97.3%; Pred. No. 2.1e-96;
Matches 365; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGGCGTCCCTGGATCGGTTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCT 60
DB 11 ATGGCGTCCCTGGATCGGTTGAAGGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCT 70
QY 61 TCGTTAGTCCATCTCTATGCCAAAATCAAGTGTGGGAAATCCATCATGAGCTGTGGGC 120
DB 71 TCGTTAGTCCATCTCTATGCCAAAATCAAGTGTGGGAAATCCATCATGAGCTGTGGGC 130
QY 121 TGCTCAGTGTGATGCAGAGTTCATGATTACAAGAGAGACCCAGAGAGAGACCTAC 180
DB 131 TGCTCAGTGTGATGCAGAGTTCATGATTACAAGAGAGACCCAGAGAGAGACCTAC 190
QY 181 TACATAGAAATATGGGATGTGGAGGCTCTGTGGGAGTGCAGCAGCGTGAAGACACA 240
DB 191 TACATAGAAATATGGGATGTGGAGGCTCTGTGGGAGTGCAGCAGCGTGAAGACACA 250
QY 241 AGAGCAGTATCTACAACTCGTAAATGGTATTATTTTCGTACAGCTTAACAAATAAG 300
DB 251 AGAGCAGTATCTACAACTCGTAAATGGTATTATTTTCGTACAGCTTAACAAATAAG 310
QY 301 AAGTCTCCCAAACTTGGCGTGTGTCATTGGAAGCTCTCAACAGGATTTGGTCCCA 360
DB 311 AAGTCTCCCAAAATTTGGCGTGTGTCATTGGAAGCTCTCAACAGGATTTGGTCCCA 370

QY 361 ACTGGAGTCTTGTTG 375
DB 371 ACTGGAGTCTTGTTG 385

RESULT 14
LOCUS CQ719549 223 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 5483 from Patent WO02068579.
ACCESSION CQ719549
VERSION CQ719549.1 GI:42280406
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
AUTHORS Kits, such as nucleic acid arrays, comprising a majority of
TITLE humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 5483 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1..223
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 31.4%; Score 223; DB 6; Length 223;
Best Local Similarity 100.0%; Pred. No. 2.4e-54;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 384 GGATTATGATCAAGAACAGTTTCTGATACCAATAACCACTGTGTGTAATAGGACTAA 443
DB 1 GGATTATGATCAAGAACAGTTTCTGATACCAATAACCACTGTGTGTAATAGGACTAA 60
QY 444 ACTGGACCAAGATTCATGAAACAAAGCGCCATGAAGTTTTAACTAGGACTGTTTCCTGGC 503
DB 61 ACTGGACCAAGATTCATGAAACAAAGCGCCATGAAGTTTTAACTAGGACTGTTTCCTGGC 120
QY 504 TGAGGATTTCAATCCAGAGAAATTAATTTGGAGTGCACAAATCCAGCTACTTAGCTGC 563
DB 121 TGAGGATTTCAATCCAGAGAAATTAATTTGGAGTGCACAAATCCAGCTACTTAGCTGC 180
QY 564 AGGTTCTTCCAATGCTGCTCAAGCTCAGTAGGTTTTTTTGATAAG 606
DB 181 AGGTTCTTCCAATGCTGCTCAAGCTCAGTAGGTTTTTTTGATAAG 223

RESULT 15
LOCUS AC133474 65838 bp DNA linear PRI 03-DEC-2002
DEFINITION Homo sapiens 3 BAC RP11-719N16 (Roswell Park Cancer Institute Human
ACCESSION AC133474
VERSION AC133474.2 GI:23334664
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 65838)
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Albrow, S.L., Amaral, H.C., Are, J.R., Ayala, M., Banks, T.,
Barbosa, J., Benton, J., Bimonte, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
```

Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Rulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Moore, S., Morgan, M., Moorish, T., Morris, S., Mosser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwotonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, O., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 65838)
Worley, K.C.

Direct Submission
Submitted (13-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 65838)
Worley, K.C.

Direct Submission
Submitted (27-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 65838)
Worley, K.C.

Direct Submission
Submitted (03-DEC-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 27, 2002 this sequence version replaced gi:22830213.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence

continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	Location/Qualifiers
source	1..65838 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosomes="3" /clone="RP11-719N16"
misc_feature	1..2005 /note="overlaps bases 117724..119726 of clone AC117472"
repeat_region	/function="clone overlap" complement(1..420) /rpt_family="L2"
STS	2133..2353 /standard_name="SHGC-68332"
repeat_region	2818..3009 /rpt_family="MIR" complement(4474..5354) /rpt_family="LIM4"
repeat_region	5358..5382 /rpt_family="(TTTTG)n" complement(5383..5642) /rpt_family="AluSx"
repeat_region	complement(5643..7192) /rpt_family="LIM4"
repeat_region	7550..7851 /rpt_family="AluY"
repeat_region	7892..8199 /rpt_family="MER1B"
repeat_region	complement(8480..8586) /rpt_family="LIMB"
repeat_region	9046..9175 /rpt_family="LIME"
repeat_region	complement(9498..10116) /rpt_family="LIMC5"
repeat_region	11023..11046 /rpt_family="AT_rich"
repeat_region	complement(12233..12522) /rpt_family="AluSg"
repeat_region	complement(12953..13127) /rpt_family="MIR"
repeat_region	complement(13247..13436) /rpt_family="MER5A"
repeat_region	13772..13813 /rpt_family="(TG)n"
repeat_region	complement(14893..15027) /rpt_family="MIR"
repeat_region	16272..16299 /rpt_family="AT_rich"
repeat_region	17561..17601 /rpt_family="AT_rich"
repeat_region	17909..18207 /rpt_family="AluY"
repeat_region	complement(18719..19081) /rpt_family="MER47A"
STS	20760..20949 /standard_name="SHGC-77253"

```

repeat_region 21591..21610
/rpt_family="TG)n"
repeat_region 22278..22315
/rpt_family="AT_rich"
repeat_region 22316..22596
/rpt_family="AluSx"
repeat_region 22597..22634
/rpt_family="(TAAA)n"
repeat_region 22957..23107
/rpt_family="AluJb"
repeat_region 23327..23425
/rpt_family="LTR29B"
repeat_region 23401..23779
/rpt_family="MER39"
repeat_region complement(24441..24627)
/rpt_family="MERSA"
repeat_region complement(25067..25188)
/rpt_family="HAL1"
repeat_region 25240..25367
/rpt_family="L2"
repeat_region 25508..25664
/rpt_family="AluJb"
repeat_region 25670..25958
/rpt_family="AluJo"
repeat_region 25959..25985
/rpt_family="AT_rich"
repeat_region complement(26771..26936)
/rpt_family="MLT1H"
repeat_region complement(27115..27217)
/rpt_family="MLT1H"
repeat_region complement(28823..29001)
/rpt_family="MIR"
repeat_region complement(29469..29599)
/rpt_family="L2"

```

```

Query Match 21.7%; Score 154.2; DB 9; Length 65838;
Best Local Similarity 98.1%; Pred. No. 5.8e-34;
Matches 156; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 383 GGGATTATGATCAAGAACAGTTTGCTGATAACCAATACCACTGTGGTAATAGGGACTA 442
Db 11832 GGGATTATGATCAAGAACAGTTTGCTGATAACCAATACCACTGTGGTAATAGGGACTA 11891

QY 443 AACTGGACCAAGTTTCATGAAACAAAGCCCATGAAGTTTAACTAGCACTGCTTCTGG 502
Db 11892 AACTGGACCAAGTTTCATGAAACAAAGCCCATGAAGTTTAACTAGCACTGCTTCTGG 11951

QY 503 CTGAGGATTTCAATCCAGAGAAATTAATTGGACTGCA 541
Db 11952 CTGAGGATTTCAATCCAGAGAAATTAATTGGATGTA 11990

```

Search completed: June 30, 2005, 23:29:36
Job time : 3261.89 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 13:44:48 ; Search time 133.331 Seconds
(without alignments)
8725.600 Million cell updates/sec

Title: US-09-945-173-3

Perfect score: 711

Sequence: 1 atgggtccctgactgggt.....agagccttcattatgactga 711

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66.4	9.3	1087	4	US-09-270-767-13908
2	41.6	5.9	1141	4	US-09-806-708B-22
c 3	36.8	5.2	7218	1	US-08-232-463-14
c 4	35.2	5.0	1141	4	US-09-806-708B-22
5	34.8	4.9	1212	2	US-09-092-770-18
6	34.8	4.9	1212	3	US-09-222-851-18
7	34.4	4.8	1182	4	US-09-248-796A-779
8	34.2	4.8	810	4	US-09-248-796A-6393
9	34	4.8	735	4	US-09-270-767-13913
10	33.8	4.8	2742	4	US-09-134-000C-2635
11	33.8	4.8	45090	4	US-09-949-016-12690
12	33.8	4.8	45091	4	US-09-949-016-16195
13	33.2	4.7	1008	4	US-09-134-000C-2379
14	33	4.6	121049	4	US-09-949-016-17513
c 15	32.8	4.6	474	4	US-09-621-976-18033
16	32.8	4.6	618	4	US-09-270-767-28849
17	32.8	4.6	709	4	US-09-270-767-12973
c 18	32.6	4.6	126254	4	US-09-949-016-15341
c 19	32.6	4.6	786431	4	US-09-751-389-3
c 20	32.4	4.6	601	4	US-09-949-016-81122
c 21	32.4	4.6	14566	4	US-09-949-016-16765
c 22	32.4	4.6	48994	4	US-09-949-016-14091
c 23	32.4	4.6	109025	4	US-09-949-016-12609
c 24	32.4	4.6	109025	4	US-09-949-016-17567
c 25	32.2	4.5	6508	4	US-09-995-317A-2
c 26	32.2	4.5	80717	4	US-09-949-016-14968
c 27	32.2	4.5	155617	4	US-09-949-016-16191

c 28	32.2	4.5	580073	4	US-08-545-528D-1	Sequence 1, Appli
c 29	32	4.5	601	4	US-09-949-016-139959	Sequence 139959,
c 30	32	4.5	148567	4	US-09-801-876B-3	Sequence 3, Appli
c 31	32	4.5	148567	4	US-10-254-869-3	Sequence 3, Appli
c 32	32	4.5	148567	4	US-10-667-442-3	Sequence 3, Appli
c 33	32	4.5	312474	4	US-09-949-016-17434	Sequence 17434, A
c 34	31.8	4.5	1164	4	US-09-710-279-3179	Sequence 3179, Ap
c 35	31.8	4.5	3603	4	US-09-710-279-3419	Sequence 3419, Ap
c 36	31.8	4.5	134890	4	US-09-949-016-15602	Sequence 15602, A
c 37	31.6	4.4	601	4	US-09-949-016-194092	Sequence 194092,
c 38	31.6	4.4	1173	4	US-09-248-796A-6337	Sequence 6337, Ap
c 39	31.6	4.4	1543	3	US-09-364-230-17	Sequence 17, Appl
c 40	31.6	4.4	17922	4	US-09-949-016-15018	Sequence 15018, A
c 41	31.6	4.4	80632	4	US-09-949-016-12951	Sequence 12951, A
c 42	31.6	4.4	137949	4	US-09-949-016-12196	Sequence 12196, A
c 43	31.6	4.4	137956	4	US-09-949-016-17260	Sequence 17260, A
c 44	31.6	4.4	202001	4	US-09-734-674-3	Sequence 3, Appli
c 45	31.4	4.4	601	4	US-09-949-016-105459	Sequence 105459,

ALIGNMENTS

RESULT 1

US-09-270-767-13908

; Sequence 13908, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 13908

; LENGTH: 1087

; TYPE: DNA

; ORGANISM: Drosophila melanogaster

US-09-270-767-13908

Query Match 9.3%; Score 66.4; DB 4; Length 1087;

Best Local Similarity 52.0%; Pred. No. 3.8e-12;

Matches 182; Conservative 0; Mismatches 156; Indels 12; Gaps 1;

Qy	1	ATGCGTCCTCGGATCGGTGAAGTACTGTGTGGGAGACTCAGGTGTGGGAAATCT	60
Db	269	ATGCCGATGAATATTCAGTTCGGAATTTGTGTGGGATTCGGTGTGGGAAGACC	328
Qy	61	TCGTTAGTCCATCTCTTATGCCAAATCAAGTGTCTGGGAAATCCATCATGGA	120
Db	329	TCCTGACGACCTGATCACCACACGAAGCCCTCATCGGCCCGCTGACGTTGGG	388
Qy	121	TGCTCAGTGTGATGTCAGATTCATGATTAAGAAGAACCCAGAGAACACCTAC	180
Db	389	TGCAACATCCAGTGAAGATGCATCCGTTGAGGAGGACCCGCTCGCGATGTCC	448
Qy	181	TACATCAATTTGGGATTTGGAGGCTCTGTGGGAGTGCACAGCGTGAAGAAGCA	240
Db	449	TTCTGTGAGCTGTTCGATGTTGGGGATTCGCTG-----AACCAAGAACACG	496
Qy	241	AGAGCAGTATTTACAACTCCGTAATATTTTGGTACACGACTTAACAAATAAG	300
Db	497	CGCAGCGTCTTCATCGGGCATCGATGGAATCATCTGTGTGACAGACCTTACCA	556
Qy	301	AAGTCTCCCAAACTTGGTGTGATTTGGAAGCTCTCAACAGGGA	350
Db	557	AAGTCTGAGGAGGAGTAAATCGACTGCTGTACGATCGTCAACAGGA	606

RESULT 2

US-09-806-708B-22

; Sequence 22, Application US/09806708B

Patent No. 6784342
GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806,708B
CURRENT FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 1141
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(1141)
OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match 5.9%; Score 41.6; DB 4; Length 1141;
Best Local Similarity 12.2%; Pred. No. 0.0017;
Matches 63; Conservative 188; Mismatches 264; Indels 1; Gaps 1;
QY 153 AGAAGCAACCCAGAGAGACCTACTACATAGATTATGGATGTTGGAGGCTCTGT 212
DB 252 MVAWTAHRRYNNGTBAYRRWTNNNNNAKACRKYWGNRABVNSTCTTWKS 311
QY 213 GGCAGTGCAGCAGCGTGAAGACACAGAGCAGTATTCTACAACTCCGTAATCGTAT 272
DB 312 KTKVTSWANNCRAGDKDKHKKWSAAMGVYNNNNNNNTYKARHBRADWVWH 371
QY 273 TATTTTCGTACAGCTTAACAATAGAAGTCTCCCAAACTGGCTGCTGGTCAAT 332
DB 372 SAWKWHANAAYSRKKWTBYKRTMVNNNGTTMKRMWAWYWKMDMBGTYNNNNNG 431
QY 333 GGAAGCTCTCAACAGGATTTGGTCCCACTGGAGTCTTGGTACAAATGGGATATGA 392
DB 432 GRTYGTWTKNKMWTYKKNCKRWDHDKTTHNTTTFWPKTYNNYKWSMTNGK 491
QY 393 TCAAGAACAGTTTGTCTGATAACCAATACCACTGTGGTAATAGGACTAACTGGACCA 452
DB 492 SHREBAAYTYWYWWRYAHANNNDYWKACTYKVBVCSKWNYYAATYKSSWY 551
QY 453 GATTCATGAACAAAGCGCATCAAGTTTAACTAGGACTGCTTTCTCGCTGAGGATTT 512
DB 552 TSRYRWKTNNSWRSDTRSMGRANNYARABHYGKWNTRWWSHTWBHBRAGAAYW 611
QY 513 CAATCCAGAAGAAATTAATTGGACTGCACAAATCCACGCTACTAGCTGCAGGTTCTTC 572
DB 612 MEMMYBAKCHMKAWYKAKYAGAGSGNNNNNNNNNNNNNNATCARDYYAASRWYAM 671
QY 573 CAATGCTGCAAGCTCAGTAGGTTTTTGTATAAGTTCATAGAGAGAGATACCTTTTAAAG 632
DB 672 ANAKWYKBAANNAYYTHANNWGCWNNATDRTFTWK-NNNNNAGTWKNNNNNAKN 730
QY 633 AGAAGTAATCAGATTCAGGCTTCTCCTGATCGGAA 668
DB 731 ASAAKNYAAAAVKAACKHWRWANKWAMRGWHAATA 766

RESULT 3

US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFELINGER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:

ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F18
US-08-232-463-14

Query Match 5.2%; Score 36.8; DB 1; Length 7218;
Best Local Similarity 3.5%; Pred. No. 0.26;
Matches 14; Conservative 210; Mismatches 172; Indels 0; Gaps 0;
QY 99 AAATCCATCATCGACTGTGGCTGCTCATGTGATGTCAGAGTTCATGATTACAAAGAAG 158
DB 1457 AAAGATAGAAGATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1398
QY 159 AACCCCAAGAGAAGACCTACTACATAGAATATAGGATGTTGGAGGCTCTGTGGCAG 218
DB 1397 RRR 1338
QY 219 TCCAGCAGCGTGAAGACAAAGACAGCAGTATTTCTACAACCTCGTAAATGGTATTATT 278
DB 1337 RRR 1278
QY 279 CGTACAGCACTTAAACAAATAAGAAGTCTCCCAAACTTCGCGTGGTTCATTGGAAGC 338
DB 1277 RRR 1218
QY 339 TCTCAACAGGATTTGGTCCCACTGGAGCTTTGGTGACAAATGGGGATTAATGATCAAGA 398
DB 1217 RRR 1158
QY 399 ACAGTTTGTGATACCAAAATACCAGTGTGGTAATAGGACTAAATGGACAGATTCATCA 458
DB 1157 RRR 1098
QY 459 TGAACAAAGCGCATGAAGTTTAACTAGGACTGC 494
DB 1097 RRR 1062

RESULT 4

US-09-806-708B-22/c

Qy 485 CTAGGACTGCTTCCCTGGTGAGATTTCAATCAGAAGAAATAAATTGGACTGCACAA 544
Db :
632 AYTGNACNGAYGGNGCNITGYWSNGARGTNGAYATHYTNAAATGGARYTNAAYATHYTA 691
Qy 545 ATCCACGGTACTTAGCTGCAGGTTCCTCCAATGCTGTCAAGCTCAGTAGGTTTTTGATA 604
Db :
692 ARGCNYNAAATGGGARYTNTGYCCNCTNACNGTNAHWSNLTGGYTNAAYTTTYYTNC 751
Qy 605 AGGTCATAGAGAAGAGATACTTTTAAAGAG 634
Db :
752 ARGTNGAYCGNTWAARGAYCTNCCNAARG 781

RESULT 7
US-09-248-796A-779
; Sequence 779, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 779
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-779

Query Match 4.8%; Score 34.4; DB 4; Length 1182;
Best Local Similarity 48.5%; Pred. No. 0.54;
Matches 95; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 381 TGGGGATATGATCATCAAGAACAGTTTGCTGATGAACCAATACCACCTGTGGTAATAGGAC 440
Db :
582 TCGAGAACAGATGAATCAATGATTAATGATTAACAACAAATGCAACGAATGCAAGCTGGACT 641
Qy 441 TAAACTGGACAGAGATTCATGAACAAAGCGCATGAAGTTTAACTAGGACTGCTTTCCT 500
Db :
642 TAAAAATATTACATGTCAGATATAAACGTCACAAATTACTTGATGAAGAGATCTTT 701
Qy 501 GGCTGAGGATTTCAATCCAGAAGAAATTAATTTGGACTGCACAAATCCACGGTACTTAGC 560
Db :
702 ATTACCGGATTATAGAATTGATTAAATTTATCGGATTTATCTAAAACCGAGTTATCAACA 761
Qy 561 TGCAGGTTCTTCCAAT 576
Db :
762 ATCAGGATTTTATGAT 777

RESULT 8
US-09-248-796A-6393
; Sequence 6393, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6393

; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2635
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2635

Query Match 4.8%; Score 33.8; DB 4; Length 2742;
Best Local Similarity 51.0%; Pred. No. 1.5;
Matches 80; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 515 ATCCAGAAAGAATTAAATTTGACTGCACAATAACCGGTACTTAGCTGACGGTTCCTTCCA 574
 ||||| ||||| ||||| ||
Db 1346 ATCCAAAAGCGTTTAAGTTAGAAACCTCATTCAGCGCCAGAGTTCGATGTCACTTACGATG 1405
 ||||| ||||| ||||| ||

QY 575 ATGCTGTCAAGCTCAGTAGGTTTTTGATAGAGTTCATAGAGAAGAGATACTTTTTAAAGAG 634
 ||||| ||||| ||||| ||
Db 1406 GTGCCGATAAATTATATGTTTTTTATGATGAAGTGAAAGAAATGACCACACTACTATCCAG 1465
 ||||| ||||| ||||| ||

QY 635 AAGGTAATCAGATCCAGGCTTTCCTGATCGAAAAAG 671
 ||||| ||||| ||||| ||
Db 1466 AGTTACTATAAGTTTGGCTTTGTTGATGAAGAGG 1502
 ||||| ||||| ||||| ||

RESULT 11
US-09-949-016-12690
; Sequence 12690, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12690
; LENGTH: 45090
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12690

Query Match 4.8%; Score 33.8; DB 4; Length 45090;
Best Local Similarity 50.3%; Pred. No. 9.6;
Matches 83; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 6 GTCCCTCGATCGGGTGAAGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTT 65
 ||||| ||||| ||||| ||
Db 2192 GACCTCCGACGGGTATATGGATGGAAGCGGACCTCGGAGCAGGTAAAGGTCCCCA 2251
 ||||| ||||| ||||| ||

QY 66 AGTCCATCTCTATGCCAAAATCAAGTCTGGGAAATCCATCATGAGCTGTGGGCTGCTC 125
 ||||| ||||| ||||| ||
Db 2252 GGAGGGGGAACGGTGCATGCTCCAAGGACTCGGGGATCCCGCATGAAAAGCGTGGTTTC 2311
 ||||| ||||| ||||| ||

QY 126 AGTGGATGTCAAGTTCATGATTACAAAGAGAACCCAGAGA 170
 ||||| ||||| ||||| ||
Db 2312 AAGTGATGGAAGCGCTCCTGATGAGGAGAGGGCTCTCCCACGA 2356
 ||||| ||||| ||||| ||

RESULT 12
US-09-949-016-16195
; Sequence 16195, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

```

; Sequence 17513, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17513
; LENGTH: 121049
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17513

Query Match      4.6%; Score 33; DB 4; Length 121049;
Best Local Similarity 51.0%; Pred. No. 35;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 384 GGATTATGATCAAGACAGTTTCTGATACCAAAATACCACTCTTGGTAATAGGGACTAA 443
Db 108039 GTATAATGATGTTAATCTTTTATTTTTCAGATCTTCAATCATTTCTTCTGTAGGGAATA 108098

QY 444 ACTGGACCATTCATGACAAACGCGCATGAGTTTAACTAGGACTGCTTCTCTGGC 503
Db 108099 AATAAAACTATTTGTATAGCTTTACCCCAATAAAATTTACCCCTTAACCTGCTCTGCTGT 108158

QY 504 TGAGGATTTCAATCCAGAGAAATTAATTGGA 536
Db 108159 AGCAGGGTAACTTCAGGAGATATATGTGA 108191

RESULT 15
US-09-621-976-18033/c
; Sequence 18033, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18033
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 16
; OTHER INFORMATION: n=a, g, c or t
; US-09-621-976-18033

Query Match      4.6%; Score 32.8; DB 4; Length 474;
Best Local Similarity 12.9%; Pred. No. 1.1;
Matches 31; Conservative 106; Mismatches 103; Indels 0; Gaps 0;

QY 291 AACAAATAGAACTCTCCAAACTTCGCTGCTGTTGTCATTTGGAAGCTCTCAACAGGGA 350
Db 308 AARRTWTARAAAYTWTWTYMKGSCMKMKMRKRYMKCCYWWWRYCCMKRKR 249
QY 351 TTTGGTGCCAACTGGAGTCTTGGTGCAAAATGGGGATTATGATCAAGAACAGTTTGCTGA 410

```

```

Db 248 GKRRARSYKGSYKGSYTYKGGKGGTGKGGKSKGRRRAMTTTTYYKKKKWSYWA 189
QY 411 TAACCAAAATACCACCTGTTGGTAATAGGGACTAAACTGGACCGAGATTCATGAAACAAAGCG 470
Db 188 ARGKTTKKWAAAMMMCMMAWYMRSMYYKKWAAAWKSCCMARKKGGSSGRWYMKRRG 129
QY 471 CCATGAAGTTTTAACTAGGACTGCTTCTCTGGCTGAGGATTTCAATCCAGAGAAATTA 530
Db 128 TCMRGRRRWYRRRRCMWTCTKGSMAAMMMSGKKRAAMCCSKYYTTCMCAAAAWWWTWA 69

Search completed: June 30, 2005, 17:01:03
Job time : 139.331 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 16:53:46 ; Search time 492.231 Seconds
(without alignments)
9059.822 Million cell updates/sec

Title: US-09-945-173-3

Perfect score: 711

Sequence: 1 atggcgtccctgagtcgggt.....agagccttcattatgactga 711

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	711	100.0	711	9	US-09-945-173-3
2	711	100.0	735	17	US-10-210-130-125
3	711	100.0	1417	13	US-10-044-090-467
4	711	100.0	1694	9	US-09-945-173-1
5	711	100.0	2398	19	US-10-479-284-34
6	707.8	99.5	739	17	US-10-210-130-127
7	617.4	86.8	959	9	US-09-764-868-321

8	444	62.4	504	10	US-09-918-995-23516	Sequence 23516, A
9	54.6	7.7	1626	20	US-10-425-115-2369	Sequence 2369, App
10	51.4	7.2	664	19	US-10-767-701-25416	Sequence 25416, A
11	50.2	7.1	1643	18	US-10-425-114-26886	Sequence 26886, A
12	50.2	7.1	1643	20	US-10-425-115-75255	Sequence 75255, A
13	45.2	6.4	761	19	US-10-437-963-650	Sequence 650, App
14	44.6	6.3	1584	19	US-10-437-963-25042	Sequence 25042, A
15	38	5.3	610	18	US-10-424-599-109350	Sequence 109350, A
16	37.4	5.3	2693	19	US-10-437-963-64985	Sequence 64985, A
17	36.4	5.1	1510	18	US-10-424-599-109349	Sequence 109349, A
18	36.2	5.1	648	21	US-10-487-901-5702	Sequence 5702, App
19	36.2	5.1	1161	19	US-10-437-963-64599	Sequence 64599, A
20	36.2	5.1	3836	19	US-10-437-963-20359	Sequence 20359, A
21	35.6	5.0	775	14	US-10-123-155-120	Sequence 120, App
22	35.6	5.0	775	15	US-10-146-731-120	Sequence 120, App
23	35.6	5.0	775	15	US-10-140-472-120	Sequence 120, App
24	35.6	5.0	775	15	US-10-141-761-120	Sequence 120, App
25	35.6	5.0	775	16	US-10-142-885-120	Sequence 120, App
26	35.6	5.0	775	16	US-10-158-790-120	Sequence 120, App
27	35.6	5.0	775	17	US-10-137-871-120	Sequence 120, App
28	35.6	5.0	775	17	US-10-140-923-120	Sequence 120, App
29	35.6	5.0	775	17	US-10-141-756-120	Sequence 120, App
30	35.6	5.0	775	17	US-10-141-759-120	Sequence 120, App
31	35.6	5.0	775	17	US-10-140-805-120	Sequence 120, App
32	35.6	5.0	775	17	US-10-140-864-120	Sequence 120, App
33	35.6	5.0	775	18	US-10-142-426-120	Sequence 120, App
34	35.4	5.0	165097	22	US-10-737-082-77	Sequence 77, Appl
35	35.4	5.0	165097	22	US-10-765-790-77	Sequence 77, Appl
36	34.8	4.9	468	14	US-10-123-155-90	Sequence 90, Appl
37	34.8	4.9	468	15	US-10-146-731-90	Sequence 90, Appl
38	34.8	4.9	468	15	US-10-140-472-90	Sequence 90, Appl
39	34.8	4.9	468	15	US-10-141-761-90	Sequence 90, Appl
40	34.8	4.9	468	16	US-10-142-885-90	Sequence 90, Appl
41	34.8	4.9	468	16	US-10-158-790-90	Sequence 90, Appl
42	34.8	4.9	468	16	US-10-063-685-150	Sequence 150, App
43	34.8	4.9	468	17	US-10-137-871-90	Sequence 90, Appl
44	34.8	4.9	468	17	US-10-140-923-90	Sequence 90, Appl
45	34.8	4.9	468	17	US-10-141-756-90	Sequence 90, Appl

ALIGNMENTS

RESULT 1
US-09-945-173-3
; Sequence 3, Application US/09945173
; Patent No. US20020127568A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 47324, A NOVEL HUMAN G-PROTEIN AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 38155-20035.00
; CURRENT APPLICATION NUMBER: US/09/945,173
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,293
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-945-173-3

Query Match 100.0%; Score 711; DB 9; Length 711;
Best Local Similarity 100.0%; Pred. No. 1.5e-213;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCGTCCCTGATCGGTGAAGTACTGGTGTGGAGACTCAGTGTGGAAATCT 60
Db 1 ATGGCGTCCCTGATCGGTGAAGTACTGGTGTGGAGACTCAGTGTGGAAATCT 60

```

QY 61 TCCTAGTCCATCTCTATGCCAAAATCAAGTCTCGGAAATCCATCATGGACTGTGGGC 120
Db 61 TCCTAGTCCATCTCTATGCCAAAATCAAGTCTCGGAAATCCATCATGGACTGTGGGC 120
QY 121 TGTCTAGTGGATGTGAGGTTTCATGATTTACAAAGAGGAAACCCAGAGAGAGACCTAC 180
Db 121 TGTCTAGTGGATGTGAGGTTTCATGATTTACAAAGAGGAAACCCAGAGAGAGACCTAC 180
QY 181 TACATAGATTTATGGATGTTGAGGCTCTGTGGCAGTGTCCAGCAGCGTGAAGACACA 240
Db 181 TACATAGATTTATGGATGTTGAGGCTCTGTGGCAGTGTCCAGCAGCGTGAAGACACA 240
QY 241 AGAGCAGTATTTACAACTCCGTAATGTTATTTTCGTACACGACTTAAACAAATAG 300
Db 241 AGAGCAGTATTTACAACTCCGTAATGTTATTTTCGTACACGACTTAAACAAATAG 300
QY 301 AAGTCTCTCCAAAATTTGGCTCTGTGTCATTTGGAAGCTCTCAACAGGAGTTTGGTGCCA 360
Db 301 AAGTCTCTCCAAAATTTGGCTCTGTGTCATTTGGAAGCTCTCAACAGGAGTTTGGTGCCA 360
QY 361 ACTGGAGTCTTGTGTAACAATGGGATTTATGATCAAGACAGTTTCTGTATTAACCAATA 420
Db 361 ACTGGAGTCTTGTGTAACAATGGGATTTATGATCAAGACAGTTTCTGTATTAACCAATA 420
QY 421 CCACTGTGTATAGGACTAACTGGACCAATTCATGAACAAAGCGCCATGAAGTT 480
Db 421 CCACTGTGTATAGGACTAACTGGACCAATTCATGAACAAAGCGCCATGAAGTT 480
QY 481 TTAATAGGACTCTTCTCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 540
Db 481 TTAATAGGACTCTTCTCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 540
QY 541 ACAATCCAGGTTACTAGTCTGAGGTTCTTCAATGCTGTCAAGCTCAGTAGTTTTT 600
Db 541 ACAATCCAGGTTACTAGTCTGAGGTTCTTCAATGCTGTCAAGCTCAGTAGTTTTT 600
QY 601 GATAAGTCTATAGAGAGACTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTTCT 660
Db 601 GATAAGTCTATAGAGAGACTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTTCT 660
QY 661 GATCGGAAAGATTTGGGCGAGGAAACATTAAGAGGCTTCATTATGACTGA 711
Db 661 GATCGGAAAGATTTGGGCGAGGAAACATTAAGAGGCTTCATTATGACTGA 711

```

RESULT 2

```

US-10-210-130-125
; Sequence 125, Application US/10210130
; Publication No. US20040014053A1
; GENERAL INFORMATION:
; APPLICANT: Zernusen, Bryan D.
; APPLICANT: Patturajan, Meera
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Berghs, Constance
; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Smithson, Glennnda
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Leite, Mario W.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Anderson, David W.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Khrantsov, Nikolai V.
; APPLICANT: Ort, Tatiana
; APPLICANT: Ellerman, Karen
; APPLICANT: Rastelli, Luca
; APPLICANT: Agee, Michele L.
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Chant, John S.
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eissen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderma, Steven K.
; APPLICANT: Hjal, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elina
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C (Cura-716 SMT)
; CURRENT APPLICATION NUMBER: US/10/210,130
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 125
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10)..(717)
; US-10-210-130-125
Query Match 100.0%; Score 711; DB 17; Length 735;
Best Local Similarity 100.0%; Pred. No. 1.5e-213;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTCCCTGGATCGGTGAAGTACTGTGTGGGAGACTCAGGTGTGGGAAATCT 60
Db 10 ATGGCGTCCCTGGATCGGTGAAGTACTGTGTGGGAGACTCAGGTGTGGGAAATCT 69
QY 61 TCGTTAGTCCATCTCTATGCCAAAATCAAGTCTCGGAAATCCATCATGGACTGTGGGC 120
Db 70 TCGTTAGTCCATCTCTATGCCAAAATCAAGTCTCGGAAATCCATCATGGACTGTGGGC 129
QY 121 TGTCTAGTGGATGTGAGGTTTCATGATTTACAAAGAGGAAACCCAGAGAGAGACCTAC 180
Db 130 TGTCTAGTGGATGTGAGGTTTCATGATTTACAAAGAGGAAACCCAGAGAGAGACCTAC 189

```

QY 181 TACATAGAAATATGGGATGTTGGAGGCTCTGTGGGAGTGGCCAGCAGCGTGAAGCACA 240
Db 190 TACATAGAAATATGGGATGTTGGAGGCTCTGTGGGAGTGGCCAGCAGCGTGAAGCACA 249
QY 241 AGAGCAGTATTTACAACTCCGTAATGGTATTAATTTTCGTACACGACTTAAACAAATAAG 300
Db 250 AGAGCAGTATTTACAACTCCGTAATGGTATTAATTTTCGTACACGACTTAAACAAATAAG 309
QY 301 AAGTCTCCCAAACTTGGCTGTTGGTCAATGGAAGCTCTCAACAGGGATTTGGTGCCA 360
Db 310 AAGTCTCCCAAACTTGGCTGTTGGTCAATGGAAGCTCTCAACAGGGATTTGGTGCCA 369
QY 361 ACTGGAGTCTTGGTGACAAATGGGATTTATGATCAAGACAGTTTGTGTAACCAATA 420
Db 370 ACTGGAGTCTTGGTGACAAATGGGATTTATGATCAAGACAGTTTGTGTAACCAATA 429
QY 421 CCACTGTTGGTAAATAGGAGCTTAACTGGACAGATTTCATGAACAAAGCGCCATGAAGTT 480
Db 430 CCACTGTTGGTAAATAGGAGCTTAACTGGACAGATTTCATGAACAAAGCGCCATGAAGTT 489
QY 481 TTAAGTCTGTTTCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 540
Db 490 TTAAGTCTGTTTCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 549
QY 541 ACAATCCAGGTAATAGCTGACAGGTTCTTCCAGTCTGTCAAGCTCAGTAGGTTTTT 600
Db 550 ACAATCCAGGTAATAGCTGACAGGTTCTTCCAGTCTGTCAAGCTCAGTAGGTTTTT 609
QY 601 GATAAGTCTATAGAGAGAGATCTTTTAAAGAGAGGTAATCAAGTTCCAGGCTTTTCT 660
Db 610 GATAAGTCTATAGAGAGAGATCTTTTAAAGAGAGGTAATCAAGTTCCAGGCTTTTCT 669
QY 661 GATCGGAAAGATTTGGGCGAGGAAACATTAAGAGCCTTCATTATGACTGA 711
Db 670 GATCGGAAAGATTTGGGCGAGGAAACATTAAGAGCCTTCATTATGACTGA 720

RESULT 3

US-10-044-090-467
; Sequence 467, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 467
; LENGTH: 1417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1307098.1
US-10-044-090-467

Query Match 100.0%; Score 711; DB 13; Length 1417;
Best Local Similarity 100.0%; Pred. No. 2.2e-213; Indels 0; Gaps 0;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGTCCCTGGATCGGTTGAAGTACTGTGTTGGGAGACTCAGGTGTTGGGAAATCT 60
Db 3 ATGCGTCCCTGGATCGGTTGAAGTACTGTGTTGGGAGACTCAGGTGTTGGGAAATCT 62
QY 61 TCGTTAGTCCATCTCTATGCAAAATCAAGTCTGGGAAATCCATCATGACTGTGGGC 120
Db 63 TCGTTAGTCCATCTCTATGCAAAATCAAGTCTGGGAAATCCATCATGACTGTGGGC 122
QY 121 TGCTCAGTGTATGTCAGAGTTTCATGATTAAGAGAGGAAACCCAGAGAGACCTAC 180
Db 123 TGCTCAGTGTATGTCAGAGTTTCATGATTAAGAGAGGAAACCCAGAGAGACCTAC 182

QY 181 TACATAGAAATATGGGATGTTGGAGGCTCTGTGGGAGTGGCCAGCAGCGTGAAGCACA 240
Db 183 TACATAGAAATATGGGATGTTGGAGGCTCTGTGGGAGTGGCCAGCAGCGTGAAGCACA 242
QY 241 AGAGCAGTATTTACAACTCCGTAATGGTATTAATTTTCGTACACGACTTAAACAAATAAG 300
Db 243 AGAGCAGTATTTACAACTCCGTAATGGTATTAATTTTCGTACACGACTTAAACAAATAAG 302
QY 301 AAGTCTCCCAAACTTGGCTGTTGGTCAATGGAAGCTCTCAACAGGGATTTGGTGCCA 360
Db 303 AAGTCTCCCAAACTTGGCTGTTGGTCAATGGAAGCTCTCAACAGGGATTTGGTGCCA 362
QY 361 ACTGGAGTCTTGGTGACAAATGGGATTTATGATCAAGACAGTTTGTGTAACCAATA 420
Db 363 ACTGGAGTCTTGGTGACAAATGGGATTTATGATCAAGACAGTTTGTGTAACCAATA 422
QY 421 CCACTGTTGGTAAATAGGAGCTTAACTGGACAGATTTCATGAACAAAGCGCCATGAAGTT 480
Db 423 CCACTGTTGGTAAATAGGAGCTTAACTGGACAGATTTCATGAACAAAGCGCCATGAAGTT 482
QY 481 TTAAGTCTGTTTCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 540
Db 483 TTAAGTCTGTTTCTGGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 542
QY 541 ACAATCCAGGTAATAGCTGACAGGTTCTTCCAGTCTGTCAAGCTCAGTAGGTTTTT 600
Db 543 ACAATCCAGGTAATAGCTGACAGGTTCTTCCAGTCTGTCAAGCTCAGTAGGTTTTT 602
QY 601 GATAAGTCTATAGAGAGAGATCTTTTAAAGAGAGGTAATCAAGTTCCAGGCTTTTCT 660
Db 603 GATAAGTCTATAGAGAGAGATCTTTTAAAGAGAGGTAATCAAGTTCCAGGCTTTTCT 662
QY 661 GATCGGAAAGATTTGGGCGAGGAAACATTAAGAGCCTTCATTATGACTGA 711
Db 663 GATCGGAAAGATTTGGGCGAGGAAACATTAAGAGCCTTCATTATGACTGA 713

RESULT 4

US-09-945-173-1
; Sequence 1, Application US/09945173
; Patent No. US20020127568A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 47324, A NOVEL HUMAN G-PROTEIN AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 38155-20035.00
; CURRENT APPLICATION NUMBER: US/09/945,173
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,293
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1694
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)...(729)
US-09-945-173-1

Query Match 100.0%; Score 711; DB 9; Length 1694;
Best Local Similarity 100.0%; Pred. No. 2.4e-213; Indels 0; Gaps 0;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGTCCCTGGATCGGTTGAAGTACTGTGTTGGGAGACTCAGGTGTTGGGAAATCT 60
Db 19 ATGCGTCCCTGGATCGGTTGAAGTACTGTGTTGGGAGACTCAGGTGTTGGGAAATCT 78
QY 61 TCGTTAGTCCATCTCTATGCAAAATCAAGTCTGGGAAATCCATCATGACTGTGGGC 120

```

Db      79  TCGTTAGTCCATCTCTCTGCCAAATCAAGTCTGGGAAATCCATCATGGACTGTGGGC 138
QY      121  TCGTCACTGATGTCAGAGTTTCATGATTACAAAGAGGAACCCCGAGAGAGACCTTAC 180
Db      139  TCGTCACTGATGTCAGAGTTTCATGATTACAAAGAGGAACCCCGAGAGAGACCTTAC 198
QY      181  TACATAGAATTTATGGAGTGTGGAGCTCTGTGGCAGTCCAGCAGCGTGAAGACACA 240
Db      199  TACATAGAATTTATGGAGTGTGGAGCTCTGTGGCAGTCCAGCAGCGTGAAGACACA 258
QY      241  AGAGCAGTATTTACAACTCCGTAATATGTTTTCGTACACAGCTTAAACAAATAAG 300
Db      259  AGAGCAGTATTTACAACTCCGTAATATGTTTTCGTACACAGCTTAAACAAATAAG 318
QY      301  AAGTCTCTCCAAAACCTTGGCTGTGGTCAATGGAGCTCTCAACAGGGAATTTGGTGCCA 360
Db      319  AAGTCTCTCCAAAACCTTGGCTGTGGTCAATGGAGCTCTCAACAGGGAATTTGGTGCCA 378
QY      361  ACTGGAGTCTTGTGCAAAATGGGATTATGATCAAGAACAGTTTCTGTATACCAATA 420
Db      379  ACTGGAGTCTTGTGCAAAATGGGATTATGATCAAGAACAGTTTCTGTATACCAATA 438
QY      421  CCACTGTGTGTAATAGGACTAACTGGACAGATTTCATGAACAAAGCGCCATGAAGTT 480
Db      439  CCACTGTGTGTAATAGGACTAACTGGACAGATTTCATGAACAAAGCGCCATGAAGTT 498
QY      481  TTAACATAGGACTCTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 540
Db      499  TTAACATAGGACTCTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 558
QY      541  ACAAATCCAGGTAATAGTCTGAGTCTTCCAAATGCTGCTCAAGCTCAGTAGGTTTTT 600
Db      559  ACAAATCCAGGTAATAGTCTGAGTCTTCCAAATGCTGCTCAAGCTCAGTAGGTTTTT 618
QY      601  GATAAGTTCATAGAGAGAGATATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTTCT 660
Db      619  GATAAGTTCATAGAGAGAGATATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTTCT 678
QY      661  GATCGGAAAGATTTGGGCGAGGAACATTAAGAGGCTTCATTATGACTGA 711
Db      679  GATCGGAAAGATTTGGGCGAGGAACATTAAGAGGCTTCATTATGACTGA 729

```

RESULT 5

```

US-10-479-284-34
; Sequence 34, Application US/10479284
; Publication No. US20040158039A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION ; YUE, Henry;
; APPLICANT: LEE, Ernestine A. ; BRCHA, Shanya D.;
; APPLICANT: BAUGHN, Mariah R. ; YAO, Monique G.;
; APPLICANT: TANG, Y. Tom ; AU-YOUNG, Janice K.;
; APPLICANT: LAL, Preeti G. ; WARREN, Bridget A.;
; APPLICANT: DUGGAN, Brendan M. ; TRAN, Uyen K.;
; APPLICANT: XU, Yuming ; THANGAVELU, Kavitha;
; APPLICANT: RICHARDSON, Thomas W. ; BANDMAN, Olga;
; APPLICANT: JONES, Karen Anne ; YANG, Junning;
; APPLICANT: EMERLING, Brooke M. ; SWARNAKAR, Anita;
; APPLICANT: LUO, Wen ; CHAWLA, Narinder K.;
; APPLICANT: AZIMZAI, Yalda ; KHAN, Farrah A.;
; APPLICANT: LU, Dnyang Aina M. ; GRIFFIN, Jennifer A.;
; APPLICANT: LEE, Soo Yeun ; BURFORD, Neill;
; APPLICANT: ELLIOTT, Vicki S. ; HONCHELL, Cynthia D.;
; APPLICANT: HE, Ann ; MASON, Patricia M.;
; APPLICANT: LI, Joana X. ; HAFALIA, April J.A.;
; APPLICANT: GURURAJAN, Rajagopal
; TITLE OF INVENTION: SECRETED PROTEINS
; FILE REFERENCE: PP-0998 USN
; CURRENT APPLICATION NUMBER: US/10/479,284
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/US02/16234
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/293,728

```

```

; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/297,019
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/299,297
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/300,537
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/301,936
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/362,439
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/363,649
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 60/366,041
; PRIOR FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PERL Program
; SEQ ID NO 34
; LENGTH: 2398
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 747515CB1
; US-10-479-284-34

Query Match      100.0%; Score 711; DB 19; Length 2398;
Best Local Similarity 100.0%; Pred. No. 2.9e-213;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGGGCGTCCCTCGATCGGCTGAAGGTACTGTGGGAGACTCAGGTGTGGGAAATCT 60
Db      3  ATGGGCGTCCCTCGATCGGCTGAAGGTACTGTGGGAGACTCAGGTGTGGGAAATCT 62
QY     61  TCGTTAGTCCATCTCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGGACTGTGGGC 120
Db     63  TCGTTAGTCCATCTCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGGACTGTGGGC 122
QY    121  TCGTCACTGATGTCAGAGTTTCATGATTACAAAGAGGAACCCCGAGAGAGACCTTAC 180
Db    123  TCGTCACTGATGTCAGAGTTTCATGATTACAAAGAGGAACCCCGAGAGAGACCTTAC 182
QY    181  TACATAGAATTTATGGAGTGTGGAGCTCTGTGGGCGAGTCCAGCAGCGTGAAGACACA 240
Db    183  TACATAGAATTTATGGAGTGTGGAGCTCTGTGGGCGAGTCCAGCAGCGTGAAGACACA 242
QY    241  AGAGCAGTATTTCTACAACTCCGTAATATGTTTTCGTACACAGCTTAAACAAATAAG 300
Db    243  AGAGCAGTATTTCTACAACTCCGTAATATGTTTTCGTACACAGCTTAAACAAATAAG 302
QY    301  AAGTCTCTCCAAAACCTTGGCTGTGGTCAATGGAGCTCTCAACAGGGAATTTGGTGCCA 360
Db    303  AAGTCTCTCCAAAACCTTGGCTGTGGTCAATGGAGCTCTCAACAGGGAATTTGGTGCCA 362
QY    361  ACTGAGTCTTGTGACAAATGGGATTTATGATCAAGAAACAGTTTCTGTATACCAATA 420
Db    363  ACTGAGTCTTGTGACAAATGGGATTTATGATCAAGAAACAGTTTCTGTATACCAATA 422
QY    421  CCACCTGTTGGTAAATAGGGACTAACTGGACCGAGATTTCATGAACAAAGCGCCATGAAGTT 480
Db    423  CCACCTGTTGGTAAATAGGGACTAACTGGACCGAGATTTCATGAACAAAGCGCCATGAAGTT 482
QY    481  TTAACATAGGACTCTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 540
Db    483  TTAACATAGGACTCTTCTCGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 542
QY    541  ACAAATCCAGGTAATAGTCTGAGTCTTCCAAATGCTGCTCAAGCTCAGTAGGTTTTT 600
Db    543  ACAAATCCAGGTAATAGTCTGAGTCTTCCAAATGCTGCTCAAGCTCAGTAGGTTTTT 602
QY    601  GATAAGGTCATAGAGAGAGATATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTTCT 660
Db    603  GATAAGGTCATAGAGAGAGATATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTTCT 662

```


; Sequence 321, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 321
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (652)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (705)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-321

Query Match 86.8%; Score 617.4; DB 9; Length 959;
Best Local Similarity 96.4%; Pred. No. 6.6e-184;
Matches 673; Conservative 1; Mismatches 18; Indels 6; Gaps 4;

QY 1 ATGGCGTCCCTGGATCGGGTGAAGGTAAGTCTGGTGTGGGAGACTCAGGTGTGGGAAATCT 60
Db |||||
QY 12 ATGGCGTCCCTGGATCGGGTGAAGGTAAGTCTGGTGTGGGAGACTCAGGTGTGGGAAATCT 71
Db |||||
QY 61 TCGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGGACTGTGGC 120
Db |||||
QY 72 TCGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGGACTGTGGC 131
Db |||||
QY 121 TCGTCACTGAGTGTGAGGTTTCAATGATTACAAAGAGGAAGACCCAGAGAGAGACTTAC 180
Db |||||
QY 132 TCGTCACTGAGTGTGAGGTTTCAATGATTACAAAGAGGAAGACCCAGAGAGAGACTTAC 191
Db |||||
QY 181 TACATAGAAATTATGGGATGTGGAGGCTCTGTGGGAGTCCAGCAGCGTGAAAGCACA 240
Db |||||
QY 192 TACATAGAAATTATGGGATGTGGAGGCTCTGTGGGAGTCCAGCAGCGTGAAAGCACA 251
Db |||||
QY 241 AGAGCAGTATTCTACAACTCCGTAATGTTATTTTGTGACACAGCTTAACAAATAG 300
Db |||||
QY 252 AGAGCAGTATTCTACAACTCCGTAATGTTATTTTGTGACACAGCTTAACAAATAG 311
Db |||||
QY 301 AAGTCTCCCAAACTTTGCTGCTGCTCAATGGAAGCTCTCAACAGGAGTTTGGTGCCA 360
Db |||||
QY 312 AAGTCTCCCAAACTTTGCTGCTGCTCAATGGAAGCTCTCAACAGGAGTTTGGTGCCA 371
Db |||||
QY 361 ACTGGAGTCTTGTGTGACAAATGGGATTTATGATCAAGAAACAGTTTCTGTATACCAAATA 420
Db |||||
QY 372 ACTGGAGTCTTGTGTGACAAATGGGATTTATGATCAAGAAACAGTTTCTGTATACCAAATA 431
Db |||||
QY 421 CCAGTCTGGTAATAGGACTAACTCGGACAGATTCATGAACAAGCGCCATGAAGTT 480
Db |||||
QY 432 CCAGTCTGGTAATAGGACTAACTCGGACAGATTCATGAACAAGCGCCATGAAGTT 491
Db |||||
QY 481 TTAACCTAGGACTGCTTTCTGCTGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 540
Db |||||
QY 492 TTAACCTAGGACTGCTTTCTGCTGCTGAGGATTTCAATCCAGAGAAATTAATTTGGACTGC 551
Db |||||
QY 541 ACAAAATCCACGGTACTTAGCTGAGGTTCTTCCAATGCTGTCAAGCTCAGTAGGTTTTT 600
Db |||||
QY 552 ACAAAATCCACGGTACTTAGCTGAGGTTCTTCCAATGCTGTCAA - CTCATAGGTTTTT 610
Db |||||
QY 601 GATAAGCTCATAGAGAGAGATCTTTTAAAGAGAGGTAATCAGATTCCAGGCTTTCT 660
Db |||||
QY 611 GAT - AGGGCATAGAGAGAGATAC - TTTTAAAGAGAGGTAATCNGATCCGAGGCTTTTG 668
Db |||||
QY 661 GATCGGAAAGATTTTGGGCGAGGAACATTTAAAGAGCCT 698

Db 669 ACTCG---AAAGATTGGCCGACCATTTAAAGCGCTT 703
|||||

RESULT 8
US-09-918-995-23516
; Sequence 23516, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23516
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(504)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23516

Query Match 62.4%; Score 444; DB 10; Length 504;
Best Local Similarity 96.4%; Pred. No. 2.9e-129;
Matches 450; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 2 TGGCGTCCCTGATCGGGTGAAGGTAAGTCTGGTGTGGGAGACTCAGGTGTGGGAAATCTT 61
Db |||||
QY 37 TTGGNANNNGGCTCGGNNGAAGGGACTGACGTTGGGAGACTCAGGTGTGGGAAATCTT 96
Db |||||
QY 62 CGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGGACTGTGGGCT 121
Db |||||
QY 97 CGTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGGACTGTGGGCT 156
Db |||||
QY 122 GCTCAGTGGATGTGAGGTTTCAATGATTACAAAGAGGAAGACCCAGAGAGAGACCTACT 181
Db |||||
QY 157 GCTCAGTGGATGTGAGGTTTCAATGATTACAAAGAGGAAGACCCAGAGAGAGACCTACT 216
Db |||||
QY 182 ACATGAAATTATGGGATGTGGAGGCTCTGTGGGAGTCCAGCAGCGTGAAAGCACA 241
Db |||||
QY 217 ACATGAAATTATGGGATGTGGAGGCTCTGTGGGAGTCCAGCAGCGTGAAAGCACA 276
Db |||||
QY 242 GAGCAGTATTCTACAACTCCGTAATGTTATTTTGTGACACGACTTAACAAATAAGA 301
Db |||||
QY 277 GAGCAGTATTCTACAACTCCGTAATGTTATTTTGTGACACGACTTAACAAATAAGA 336
Db |||||
QY 302 AGTCTCTCCCAAACTTTGCTGCTGCTCAATGGAGCTCTCAACAGGAGTTTGGTGCCAA 361
Db |||||
QY 337 AGTCTCTCCCAAACTTTGCTGCTGCTCAATGGAGCTCTCAACAGGAGTTTGGTGCCAA 396
Db |||||
QY 362 CTGGAGTCTTGTGTGACAAATGGGATTTATGATCAAGAAACAGTTTGTGTATACCAAATAC 421
Db |||||
QY 397 CTGGAGTCTTGTGTGACAAATGGGATTTATGATCAAGAAACAGTTTGTGTATACCAAATAC 456
Db |||||
QY 422 CACTGTTGTAATAGGACTTAACTGGACCAAGTTCATCAACAAAG 468
Db |||||
QY 457 CACTGTTGTAATAGGACTTAACTGGACCAAGTTCATCAACAAAG 503
Db |||||

RESULT 9
US-10-425-115-2369
; Sequence 2369, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 2369
; LENGTH: 1626
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_102161C.1
US-10-425-115-2369

Query Match 7.7%; Score 54.6; DB 20; Length 1626;
Best Local Similarity 60.4%; Pred. No. 5.8e-06;
Matches 90; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 3 GCGCTCCCTGGATCGGGTGAAGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTC 62
Db 342 GCGGCCCTGGCGCCAGGTCGTCTACTCGTGTGGGACTCAGGTGTGGGAAATCTTC 401
QY 63 GTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGACTGTGGGCTG 122
Db 402 ATTGGTGCATCTCTTACTGAAAGATTCTGCAGTGGCTCGACCCAGCCCAAAATGGGTG 461
QY 123 CTCAGTGGATGTCAGAGTTCATGATTACA 151
Db 462 TGCAGTGGTGTAAACATATTACTTACA 490

RESULT 10

US-10-767-701-25416
; Sequence 25416, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 25416
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30951803
US-10-767-701-25416

Query Match 7.2%; Score 51.4; DB 19; Length 664;
Best Local Similarity 59.1%; Pred. No. 3.7e-05;
Matches 88; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 3 GCGCTCCCTGGATCGGGTGAAGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTC 62
Db 398 GCGGCCCTGGCGCCAGGTCGTCTACTCGTGTGGGACTCAGGTGTGGGAAATCTTC 457
QY 63 GTTAGTCCATCTCTATGCCAAATCAAGTCTGGGAAATCCATCATGACTGTGGGCTG 122
Db 458 ATTGGTGCATCTCTTATTGAAAGTCTTGCAGTGGCTCGACCCAGCCCAAAATGGGTG 517
QY 123 CTCAGTGGATGTCAGAGTTCATGATTACA 151
Db 518 TGCAGTGGTGTAAACATATCACTTACA 546

RESULT 11

US-10-425-114-26886
; Sequence 26886, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26886
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4666-008-F5_FLI
US-10-425-114-26886

Query Match 7.1%; Score 50.2; DB 18; Length 1643;
Best Local Similarity 59.4%; Pred. No. 0.00014;
Matches 85; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 8 CCCTGGATCGGGTGAAGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTTAG 67
Db 325 CCTTCGGCAAGTGGCGTCCTCATCGTCGGTATTCAGGTGTGGGAAATCTTCATGG 384
QY 68 TCCATCTCCTATGCCAAATCAAGTCTGGGAAATCCATCATGACTGTGGGCTGCTCAG 127
Db 385 TGCATCTCATTTGAAAAGTCTGCCATTGCTCGACCATCTCAACAGTAGGATGCACTG 444
QY 128 TGGATGTCAGAGTTCATGATTAC 150
Db 445 TGGGCATTAAACATGTTACTTAC 467

RESULT 12

US-10-425-115-75255
; Sequence 75255, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 75255
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_168664C.1
US-10-425-115-75255

Query Match 7.1%; Score 50.2; DB 20; Length 1643;
Best Local Similarity 59.4%; Pred. No. 0.00014;
Matches 85; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 8 CCCTGGATCGGGTGAAGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTTAG 67
Db 325 CCTTCGGCAAGTGGCGTCCTCATCGTCGGTATTCAGGTGTGGGAAATCTTCATGG 384
QY 68 TCCATCTCCTATGCCAAATCAAGTCTGGGAAATCCATCATGACTGTGGGCTGCTCAG 127

Db 385 TGCATCTCATTTTGAAGAAGTTCTGCCATTGCTCGACCATCTCAACACAGTAGGATGCACTG 444
QY 128 TGGATGTCAGAGTTCATGATTAC 150
Db 445 TGGCCATTAAACATGTTACTTAC 467

RESULT 13

US-10-437-963-650
; Sequence 650, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 650
; LENGTH: 761
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(761)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_100593C.1
US-10-437-963-650

Query Match 6.4%; Score 45.2; DB 19; Length 761;
Best Local Similarity 50.0%; Pred. No. 0.0036;
Matches 175; Conservative 0; Mismatches 163; Indels 12; Gaps 2;
QY 2 TGGCGTCCCTGGATCGGTGAAGTACTGTGTGGGAGACTCAGGTGTGGGAAATCTT 61
Db 122 TGGGCCCTTCGCCCAAGTCGCGTCTCTGTCGCGGATTCAGGTGTGGGAAATCTT 181
QY 62 CGTTAGTCCATCTCTATGCCAAATCAAGTGTCTGGGAAATCCATCATGGAAGTGTGGGCT 121
Db 182 CTTTGTGTCATCTCATTTTAAAGGTTCTGCTATTGCTGACCTGCCAGACAGTAGGAT 241
QY 122 GCTCAGTGAATGCAGAGTTCATGATTACAAAGAGAACCCAGAGAGAGAGACTACT 181
Db 242 GTGCGGTTGGCGTTAAACATATTACTATTGGAAGTGCAGCGGTTCTTCTATAACATCA 301
QY 182 ACATAGAAATATGGGA-----TGTCGAGCTCTGTGGGAGTGCAGCA-----GCG 229
Db 302 TCAGCCAGCTTCAAAGGAACCTCTTTGTGAGCTTTGGGATGTTTCAGGACATGAACGCT 361
QY 230 TGAAGACACAGAGCAGATTTCTCAACACTCCGTAATGTTATTTTCGTACAGACT 289
Db 362 ACAGACATGCCGTTCAATTTTCTATACACAAATTAATGTTGTCATATTGTTATGACC 421
QY 290 TACAAATAGAAAGTCCCTCCAAAACCTTGGTGTGGTTCATTTGGGAAGCT 339
Db 422 TCTCTCAGAGGAGACCAAAACAAATTTGAATAAATGGGAGTTGAAGTT 471

RESULT 14

US-10-437-963-25042/c
; Sequence 25042, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 25042
; LENGTH: 1584
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29969C.1
US-10-437-963-25042

Query Match 6.3%; Score 44.6; DB 19; Length 1584;
Best Local Similarity 57.6%; Pred. No. 0.0082;
Matches 80; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 16 CGGGTGAAGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTAGTCCATCTC 75
Db 1223 CAGGTGCGGTCTGCTGTCGCGGACTCAGGTGTGGGAAATCATCATTTGGTTCATCTC 1164
QY 76 CTATGCCAAATCAAGTGTCTGGGAAATCCATCATGGAATGCTGGCTGTGCTCAGTGGATGTC 135
Db 1163 ATTCTGAAGGCTCTGCAATTGCTGCACCAACCCCAACGTTGGATGCGCAGTTGATGT 1104
QY 136 AGAGTTCATGATTACAAAG 154
Db 1103 AAATATATTTTATTGAAG 1085

RESULT 15

US-10-424-599-109350
; Sequence 109350, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 109350
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(610)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_69758C.1
US-10-424-599-109350

Query Match 5.3%; Score 38; DB 18; Length 610;
Best Local Similarity 55.2%; Pred. No. 0.6;
Matches 74; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 16 CGGGTGAAGTACTGGTGTGGGAGACTCAGGTGTGGGAAATCTTCGTAGTCCATCTC 75
Db 353 CAGGTCAAGATCTCTCGTTGTTGTTGATTCAGGTGTGGGAAAGACTTCTTTAGTTAACCTG 412

Qy	76	CTATGCCAAATCAAGTGTGCGGAAATCCATCATGGAAGTGTGGGCTGCTCAGTGGATGTC	135
Db	413	ATTGTAAAGGTTCTTCAATTGCTCGCCCTCTCAACAATGGTTGTTTCAGTTGATGTG	472
Qy	136	AGAGTTTCATGATTA	149
Db	473	AGCATATTACTTA	486

Search completed: June 30, 2005, 23:57:30
 Job time : 495.231 secs

THIS PAGE BLANK (USPTO)